

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number
WO 2004/046376 A2

(51) International Patent Classification⁷: **C12Q 1/68**,
C12N 15/12, C07K 14/47, 16/18, G01N 33/68, A61K
31/712, 31/713, 39/00

(DK). LEFFERS, Henrik [DK/DK]; Skt. Hans Gade
28, st., DK-2200 Copenhagen N (DK). KJEMS, Jørgen
[DK/DK]; Solbakken 15, DK-8240 Risskov (DK).

(21) International Application Number:
PCT/DK2003/000794

(74) Agent: PLOUGMANN & VINGTOFT A/S; Sundkrogs-
gade 9, Post Office Box 831, DK-2100 Copenhagen Ø
(DK).

(22) International Filing Date:
19 November 2003 (19.11.2003)

(81) Designated States (*national*): AE, AG, AL, AM, AT (uti-
lity model), AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ,
CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE
(utility model), DE, DK (utility model), DK, DM, DZ, EC,
EE (utility model), EE, EG, ES, FI (utility model), FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,
MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL,
PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL,
SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN,
YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2002 01792 19 November 2002 (19.11.2002) DK

(71) Applicants (*for all designated States except US*):
H:S RIGSHOSPITALET [DK/DK]; Blegdamsvej 9,
DK-2100 Copenhagen Ø (DK). ÅRHUS UNIVERSITET
[DK/DK]; Ndr. Ringgade 1, DK-8000 Århus C (DK).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): HERTZ, Anne,
Mette, Buhl [DK/DK]; Gl. Byvej 17, DK-2979 Hørsholm

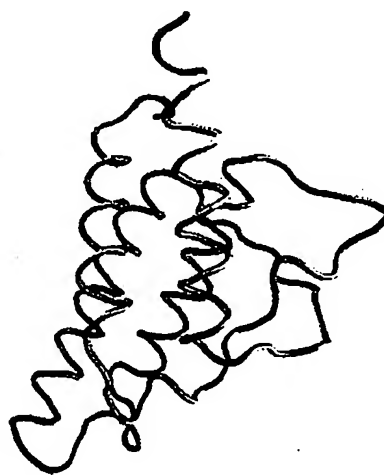
(84) Designated States (*regional*): ARIPO patent (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,

[Continued on next page]

(54) Title: METHODS AND KITS FOR DIAGNOSING AND TREATING B-CELL CHRONIC LYMPHOCYTIC LEUKEMIA
(B-CLL)



AMB1



Human IL4

(57) Abstract: The present invention relates to methods and kits for detecting several polynucleotide sequence found to be indicative of a poor prognosis of B-CLL. All the polynucleotides are transcribed from a region on human chromosome 12p21-22. Most of the polynucleotides do not encode larger polypeptides, but may encode small peptides, they may function as RNAs. Four polynucleotides encode a novel protein, which in one preferred embodiment can be used as a cytokine, preferably as an interleukin. Furthermore the invention relates to methods and compositions for treating B-CLL in particular poor prognosis B-CLL.



SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- *without international search report and to be republished upon receipt of that report*

Methods and kits for diagnosing and treating B-Cell Chronic Lymphocytic Leukemia (B-CLL)

All patent and non-patent references cited in the present application, are hereby incorporated by reference in their entirety.

5 Field of invention

The present invention relates to methods and kits for detecting several polynucleotide sequence found to be indicative of a poor prognosis of B-CLL. All the polynucleotides are transcribed from a region on human chromosome 12p21-22. Most of the polynucleotides do not encode larger polypeptides, but may encode small peptides, they may function as
10 RNAs. Four polynucleotides encode a novel protein, which in one preferred embodiment can be used as a cytokine, preferably as an interleukin. Furthermore the invention relates to methods and compositions for treating B-CLL in particular poor prognosis B-CLL.

Background of invention

B-CLL is the most common form of leukaemia in Denmark, with more than 250 new cases
15 diagnosed every year. The disease results in accumulation of CD19+CD5+CD23+ lymphocytes in the blood, bone marrow and organs of the patients. B-CLL cells are long-lived, slowly dividing and locked in the G₁ phase of the cell cycle. At this time it is unknown how or why B-CLL occurs and no cure is known for B-CLL. The application of more aggressive treatment strategies has been hampered by the inability to identify
20 reproducible and reliable prognostic predictors in patients with poor outcome in this disease. In many patients the diagnosis does not affect morbidity or mortality. Other patients suffer from an incurable cancer that inevitably results in death, regardless of treatment. Until recently this latter group of patients could not be identified at the time of diagnosis. Recently, two studies established the mutational status of immunoglobulin
25 variable region of the heavy chain (Ig V_H) genes in B-CLL as independent prognostic markers, within each clinical stage (Damle, et al. & Hamblin, et al.). Patients without somatic hypermutation show much shorter survival than patients with somatic hypermutation. FISH-studies of cytogenetic aberrations in B-CLL established specific abnormalities on chromosomes 11 (ATM), 12 (?), 13 (Leu-1 and-2) and 17 (p53) as
30 independent prognostic markers, within each clinical stage (Dohner, et al.). Very recent studies have demonstrated that independent risk prediction, using a combined analysis of Ig V_H gene mutational analysis and cytogenetics, can identify subgroups of B-CLL with median survivals ranging from less than 2.5 years to more than 15 years (Krober, et al., Lin, et al., & Oscier, et al.) (see Figure 1). Since the process of characterising the Ig V_H
35 gene mutational status of an individual patient is cumbersome, it is desirable to provide easier tests based on diagnostic markers for use in the differential diagnosis of such cancer patients.

Summary of invention

It is an object of preferred embodiments of the present invention to provide differentially expressed transcription products, which can be used as prognostic markers of disease and give information about the differences in etiology between different groups of B-CLL patients. These differentially expressed transcription products are genetic markers that can be used in an easy assay to distinguish between subgroups of B-CLL patients and especially identify B-CLL patients with a poor prognosis.

This method for diagnosing a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) comprises the steps of determining the presence or absence of at least one expression product such as a transcriptional product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from a individual. As evidenced by the appended examples, the present inventors have determined that the expression products of this invention are present in one subtype of B-CLL having poor prognosis and thus of great diagnostic value and independent prognostic value. Equally important, an expression product comprising a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 has not been found in any of the other tissue types tested (see e.g. Figure 8).

The vast majority of patients which show expression of the AMB-1 gene in form of at least one of the expression products selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 show unmutated Ig V(H) genes which is consistent with poor prognosis B-CLL. The presence of an expression product of the AMB-1 gene can be determined easily using standard laboratory procedures and equipment. Therefore the diagnostic method provided by the present inventors provides an easy method of diagnosis as compared to the determination of the mutation status of Ig V(H) genes and can furthermore give additional information about the prognosis.

Accordingly, a further object of preferred embodiments of the present invention is a method for determining the progress of B-CLL comprising determining the amount of at least one expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual. The method may be used e.g. for determining the efficiency of a treatment, i.e. to see whether the amount of the expression product decreases or increases in response to a curative treatment.

The expression products of the present invention are all transcripts of SEQ ID NO:1 and/or SEQ ID No:5, the gene of the present invention called AMB-1 which also encodes a novel polypeptide (SEQ ID NO:3).

A further object of preferred embodiments of the present invention is to provide a cure and/or treatment of patients with B-CLL, in particular of patients with poor prognosis B-CLL such as the sub-type of B-CLL which is characterised by the presence of an expression product of the present invention.

5

The method for treating B-CLL comprises administering to an individual with a B-CLL diagnosis a compound capable of decreasing or inhibiting the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5. This expression product preferably comprises a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. The present inventors believe that the presence of at least one of said expression products is an etiological factor in B-CLL and that the disease can be treated or cured by inhibiting the expression of at least one of such products and/or by inhibiting the effect of such product by e.g. rendering it inactive.

15

A further preferred object of embodiments of the present invention is to destroy or to eliminate the transcription of at least one expression product comprising at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 for the treatment of cancer, such as a poor prognosis sub-type of B-CLL.

20

The above destruction or elimination is obtained by applying polynucleotides or oligonucleotides in the form of small interfering RNA molecules (siRNA), antisense molecules or ribozymes.

25

In one aspect the invention relates to a gene therapy vector capable of inhibiting or decreasing the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5, said gene therapy vector preferably encoding a specific siRNA molecule, a specific antisense molecule or a specific ribozyme being capable of decreasing or inhibiting the formation of an expression product of SEQ ID NO:1 and/or SEQ ID NO:5. This gene therapy vector can be used for treating B-CLL based on the finding that the AMB-1 gene encoded by SEQ ID No:1 and/or SEQ ID No:5 is an etiological factor in B-CLL.

30

Both SEQ ID No 1 which is a 20,000 nucleotide long sequence and SEQ ID No 5 which is a 80,000 nucleotide long sequence provides several transcriptional products in B-CLL cells in patients with poor prognosis B-CLL. Some of the transcriptional products e.g. SEQ ID No 2 and SEQ ID No 4 consists of two exons (SEQ ID No: 15 and SEQ ID No: 16) separated by the same intron. Both mRNA sequences encode an open reading frame (SEQ ID No: 17) encoding a 121 amino acid peptide (SEQ ID No 3).

40

Accordingly, yet another object of preferred embodiments of the present invention relates to a novel class of polypeptides. These may be described as a group of isolated polypeptides or proteins comprising or essentially consisting of the amino acid sequence of SEQ ID No. 3, or a fragment thereof, or a polypeptide functionally equivalent to SEQ ID

No. 3, or a fragment thereof, wherein said fragment or functionally equivalent polypeptide has at least 60% sequence identity with the polypeptide of SEQ ID No 3. The polypeptides of the present invention may have interleukin or cytokine activity.

5 In a still further aspect the invention relates to an isolated polynucleotide selected from the group consisting of:

- i) a polynucleotide comprising nucleotides of SEQ ID No 5,
- ii) a polynucleotide encoding a polypeptide having the amino acid sequence of SEQ ID No
- 10 3,
- iii) a polynucleotide, the complementary strand of which hybridises, under stringent conditions, with a polynucleotide as defined in any of i) and ii).
- iv) a polynucleotide which is degenerate to the polynucleotide of iii), and
- v) the complementary strand of any such polynucleotide.

15

The polypeptides encoded by the polynucleotides may furthermore

- a) have at least 60 % sequence identity with the amino acid sequence of SEQ ID No 3 and have interleukin or cytokine activity,
- b) be recognised by an antibody, or a binding fragment thereof, which is capable of
- 20 recognising an epitope, wherein said epitope is comprised within a polypeptide having the amino acid sequence of SEQ ID No 3; and/or
- c) be competing with a polypeptide having the amino acid sequence as shown in SEQ ID No 3 for binding to at least one predetermined binding partner such as a cytokine receptor.

25

One further therapeutic application of the present invention is a method of vaccination against B-CLL said method comprising immunising an individual against a translational product of SEQ ID No:1 and/or SEQ ID No:5. By stimulating the immune system of an individual to produce antibodies against the translational product the individual can

30 become immune towards B-CLL and/or the method can be used as part of therapy. The state of the art describes various ways of immunising an individual against a particular protein.

Finally, the invention provides a method for determining an increased or decreased

35 predisposition for B-CLL comprising determining in a biological sample from an individual a germline alteration in a target nucleic acid sequence comprising 150,000 nucleotides, said target nucleic acid sequence comprising at least 10 nucleotides of SEQ ID No:1 and/or SEQ ID No:5. This aspect is based on the finding of the importance of the expression product of SEQ ID No:1 and/or SEQ ID No:5, and the absence of any detectable expression product

40 of SEQ ID No:1 and/or SEQ ID No:5 in healthy tissue and in patients with good prognosis B-CLL. It is highly likely that the difference is caused by a germline alteration. A germline alteration can be targeted by gene therapy methods and by the methods provided in the present invention.

Description of Drawings

Figure 1: Overall survival of B-CLL patients by genotype (all stages) The prognostic significance of V_H homology and cytogenetic aberrations is independent of clinical stage (from Kröber et al., 2002 (4)).

- 5 Figure 2 (a-d): Survival curves for survival or time to progression based on *AMB-1* expression or IgV_H mutational status respectively. Patients are newly diagnosed, untreated B-CLL patients (n=34).

Figure 3: Northern blotting demonstrating the expression level of *AMB-1* in various tissue samples using as a probe an 875 base pair fragment of Exon 3/Seq ID No:16. UPN 1, 4 and 7 are unmutated B-CLL patients, UPN 19, 9, 10, 13 and 21 are mutated B-CLL patients. Included are RNA samples from normal Colon, Spleen, Bone Marrow and PBL (peripheral blood lymphocytes) and RNA from the Ramos and Granta cell lines. Equal loading of lanes was confirmed by re-probing with an actin probe (results not shown).

15

Figure 4. Alignment of *AMB1* with *IL4* based on structural similarity. *IL4* is called *d1lara* in the alignment. The additional lines indicate the structural similarity.

Figure 5. A 3D search, where the peptide sequence has been searched for similarity to
20 known protein or peptide 3D-structures.

Figure 6. Predicted 3-D structure of *AMB-1* compared to the known 3-D structure of human *IL4*. Prediction is performed using SEQ ID No:3 and the method described in: Enhanced Genome Annotation using Structural Profiles in the Program 3D-PSSM. Kelley LA,
25 MacCallum RM & Sternberg MJE (2000). J. Mol. Biol. 299(2), 499-520.

Figure 7. Alignment of the *AMB1* peptide sequence with the sequences of *IL4*, *IL3*, *IL13* and GM-CSF, based on their structures.

30 Figure 8. A table showing the tissue types on the MTE array used for dot blotting of *AMB-1* to check for expression in other tissue types.

Figure 9. Schematic representation of the transcriptional products of the present invention compared to the genomic *AMB1* sequence (1 & 5). 2 is mRNA short form (SEQ ID No 2). 4
35 is mRNA long form (SEQ ID No 4). 6 (SEQ ID No:6), 7 (SEQ ID No:7), 8 (SEQ ID No:8), 9 (SEQ ID No:9), 10 (SEQ ID No:10) and 11 (SEQ ID No:11) are alternative transcription products all comprising at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

40

Figure 10. RT-PCR demonstrating the expression of *AMB-1* in B-CLL patients by RT-PCR. UPN1-UPN8 are unmutated patients, UPN9-UPN16 are mutated patients.

Detailed description of the invention

The present invention in particular relates to polynucleotide sequences found to be indicative of a poor prognosis of B-CLL and new methods and compositions for treating B-CLL in particular poor prognosis of B-CLL. An overview of the sequences disclosed by the present invention is present here:

SEQ ID No:1 includes 20.000 bp human genome sequence, derived from BAC clone acc. no. AC063949. It includes the region encoding the mRNAs corresponding to the cDNAs described in Seq ID Nos:2, 4, 6, 7, 8, 9, 10, and 11 and possible up- and down-stream regulatory sequences. Seq ID No:1 includes a subset of the sequence described in Seq ID No:5. The sequence is derived from human 12q21-22.

SEQ ID No:2 includes a cDNA corresponding to a putative mRNA transcript that includes the region encoding the peptide sequence in Seq ID No:3 and up- and downstream regions. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22. From position 2317, this sequence is identical to Seq ID No:4.

Seq ID No:3 contains a peptide sequence encoded by some mRNAs transcribed from the region on human chromosome 12q21-22 included in Seq ID No:1 and Seq ID No:5. It is encoded by the mRNA sequences identified as cDNAs in Seq ID Nos 02, 04, 09 and 11.

Seq ID No:4 includes a cDNA corresponding to a putative mRNA transcript that includes the region encoding the peptide sequence in Seq ID No:3 and up- and downstream regions. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.

Seq ID No:5 includes 80.000 bp human genome sequence, derived from BAC clone acc. no. AC063949. It is an expansion of the genomic sequence included in Seq ID No:1. which is contained within this sequence. It includes the region encoding the mRNAs corresponding to the cDNAs described in Seq ID Nos 02, 04, 06, 07, 08, 09, 10 and 11 and possible up- and down-stream regulatory sequences. The sequence is derived from human 12q21-22.

Seq ID No:6 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes two exons. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.

Seq ID No:7 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes three exons, the first and third are identical to the two exons in Seq ID No:6. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5.doc from human 12q21-22.

Seq ID No:8 corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes two exons, the last is also present as exon-2 in a human cDNA clone

(sequence acc. no. BC036936) (Seq ID No:9.doc). It is transcribed from the +strand of Seq ID No:5.doc from human 12q21-22.

- Seq ID No:9** corresponds to a human cDNA sequence (sequence acc. no. BC036936). It is transcribed from the +strand Seq ID No:5 from 12q21-22. We have not cDNA cloned this cDNA, but a splice variant (Seq ID No:8), where exon-2 of this sequence was spliced to exon-1 of Seq ID No:6 was detected by cDNA cloning. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.
- 10 **Seq ID No:10** corresponds to a cDNA detected partly by cDNA cloning, partly by PCR analysis, corresponding to an mRNA transcript. It includes two exons, exon-1 includes the region encoding Seq ID No:3 and exon-1 from Seq ID No:6; exon-2 is identical to exon-2 in Seq ID No:6 and exon-3 in Seq ID No:7. It is transcribed from the +strand of Seq ID No:1.doc and Seq ID No:5 from human 12q21-22.
- 15 **Seq ID No:11** corresponds to a cDNA detected by cDNA cloning, corresponding to an mRNA transcript. It includes one exon. The sequence includes the the region that encodes Seq ID No:3, exons 2 and 3 from Seq ID No:7 and the region between those exons. It is transcribed from the +strand of Seq ID No:1 and Seq ID No:5 from human 12q21-22.
- 20 **Seq ID No:12** Is an exon sequence. It corresponds to the first exon in Seq ID No: 9 It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.
- Seq ID No:13** Is an exon sequence. It corresponds to the first exon in Seq ID No: 6, 7, 8, and it is included in Seq ID No: 2. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.
- 25 **Seq ID No:14** Is an exon sequence. It is identical to Seq ID No:13, but with an additional GT dinucleotide at the 3'end, caused by the use of an alternative splice site. It can replace Seq ID No:13 as the first exon in Seq ID No: 6, 7, 8, and be included in Seq ID No: 2. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.
- 30 **Seq ID No:15** Is an exon sequence. It corresponds to the second exon in Seq ID No:7 and it is included in Seq ID No: 2, 4, 10 and 11. It is transcribed from the +strand of Seq ID No:1 and 5 from human 12q21-22.
- 35 **Seq ID No:16** Is an exon sequence. It corresponds to the third exon in Seq ID No:7, it is the second exon in Seq ID No:6 and 11 and it is included in Seq ID No:2 and 4. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.
- 40 **Seq ID No:17** Is the sequence encoding the peptide in Seq ID No: 3. It is included in Seq ID No:2, 4, 10 and 11. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

Seq ID No:18 Is an exon sequence. It corresponds to the second exon in Seq ID No:8 and 9. It is transcribed from the +strand of Seq ID No:5 from human 12q21-22.

5 Methods of diagnosis

One important aspect of the present invention relates to diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) having poor prognosis. A further important aspect of the invention relates to prognosis of B-CLL. These methods are based on the discovery by the present inventors that an expression product which comprise at least one nucleotide
10 sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 is (are) only present in particular subtypes of B-CLL associated with poor prognosis and completely absent in other subtypes of B-CLL and in healthy tissue (see in particular example 2). By completely absent is meant that the expression products are not detected in any of the
15 other tissue types with the methods used in the appended examples. This is indicative of a complete absence of any transcript or a very low level of transcript in the other tissue types.

The expression product is encoded by SEQ ID No 1 and/or SEQ ID No 5, and the
20 expression product is selected from the group consisting of transcriptional products and translational products.

Thus, the present invention relates to a method for detecting the presence or absence of at least one expression product, wherein the at least one expression product comprise a
25 nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

×

30 "Expression product" is herein meant to be a product which is the result of the expression of a polynucleotide such as DNA sequence, e.g. a genomic DNA sequence, and is in the form of either a polypeptid or in the form of a polynucleotide, i.e. an expression product can be selected from the group consisting of a transcriptional product and a translational product. In the case where the expression product is a polynucleotide, said polynucleotide
35 is preferably mRNA selected from the group consisting of mRNA, pre-mRNA, pre-pro-mRNA.

A "transcriptional product" or a "transcription product" is herein meant to be a product resulting from a transcription of a polynucleotide such as a DNA molecule, preferably a
40 genomic DNA molecule. A transcriptional product is inherently a nucleotide, such as an oligonucleotide or a polynucleotide.

A "translational product" or a "translation product" is herein meant to be a product resulting from a translation of a transcriptional product such as a mRNA. A translational product is inherently a oligopeptide or a polypeptide.

- 5 The expression product of the present invention has almost exclusively been found as transcription products in patients with poor B-CLL prognosis. Based on the experimental data presented in the herein, the inventors expect that it turns out that the subtype of B-CLL is characterised solely or better by the presence of a transcriptional or translational product which comprise a sequence selected from the group consisting of SEQ ID No:3,
10 SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

- Preferably the individual is a mammal, more preferably a human. It is also expected that the gene encoded by SEQ ID No 1 and/or SEQ ID No 5 and the expression products
15 derived from said gene can be used as a diagnostic tool in other species in particular in mammals selected from the group: domestic animals such as cow, horse, sheep, pig; and pets such as cat or dog.

- In the case that the expression product is a transcriptional product, this transcriptional
20 product just needs to comprise at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. These transcriptional products will preferably be derived from SEQ ID No 1 and/or SEQ ID No 5 and may be in the form of mRNA or any pre- or pro-forms of said mRNA. As described, the transcriptional product may comprise at
25 least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as two of the nucleotide sequences selected from the group consisting
30 of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, such as 3 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, e.g. 4 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID
35 No:16, SEQ ID No:17 and SEQ ID No:18, such as 5 of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 and ultimately the transcriptional product may comprise all of the six nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17
40 and SEQ ID No:18. The transcriptional product of the present invention can have any sequence which is a result of combining the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 as long as the specific nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15,

SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 can be identified as intact sequences in the transcriptional product.

Examples of transcriptional products in the form of specific mRNAs which comprise at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 are the sequences corresponding to SEQ ID No 2 (short cDNA clone) SEQ ID No 4 (long cDNA clone) SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11. These mRNA sequences have been found in patients with poor prognosis.

10

It is obvious for a person skilled in the art that any fragments of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 will have the same diagnostic value as long as the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 can be found in the fragments.

15

The mRNA sequence may be detected in a sample using hybridisation techniques. In particular when more than one analysis is to be performed at the same time it is advantageous to use a DNA array comprising e.g. an oligomer of at least 15 consecutive bases selected from the group consisting of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10, SEQ ID No:11, SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

Another way of detecting the presence or absence of the transcriptional product is by specifically amplifying the transcriptional product having a sequence corresponding to SEQ ID No 2, 4, 6, 7, 8, 9, 10 or 11 or fragments thereof. This can be done by selecting primer pairs which cause only the amplification of these sequences.

25

Generally, hybridisation techniques are selected from not limited to the group consisting of in situ hybridisation, northern blots, Southern blots, dot blots and PCR based techniques.

30

A non-limiting list of PCR based techniques include rt-PCR, quantitative PCR and realtime PCR.

According to yet another embodiment, the translational product is a protein encoded by a polynucleotide selected from the group consisting of SEQ ID No:1, SEQ ID No:5, SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11. Detection of this protein can be done with state of the art methods including the detection with an antibody directed against said protein, such as Western blotting, more preferably by using a fluorescently labelled antibody, preferably wherein the method comprises the use of flowcytometry, such as FACS. Other methods include but are not limited to gel electrophoresis, gel filtration, ion exchange chromatography, FPLC, mass spectrometry and immunohistochemistry.

40

Preferably, said protein is selected from the group comprising SEQ ID No 3 (protein), or a protein sharing at least 60 % sequence identity with SEQ ID No 3. The protein with the amino acid sequence set forth in SEQ ID No 3 is the longest open reading frame in the cDNA sequence of SEQ ID No 2 or 4.

5

In a specific embodiment of the present invention is a method for determining whether an individual has a B-CLL sub-type with poor prognosis, the method comprising determining the level of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 of said individual, and indicating the individual as having a B-CLL sub-type with poor prognosis if the level of the expression product is at or beyond a discriminating value and indicating the individual as not having a B-CLL sub-type with poor prognosis if the level of the expression product is not at or beyond the discriminating value, the discriminating value being a value which has been determined by measuring the level of the expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in both a healthy control population and a population with known B-CLL sub-type with poor prognosis, thereby determining said discriminating value which identifies the B-CLL sub-type population having a poor prognosis.

25

In this method the individual may be a member of an unselected population or be a member of a population already identified as having a B-CLL sub-type with a poor prognosis.

25

The above method may be performed such that the determination is performed at several time points at intervals as part of a monitoring of a cancer patient after or during the treatment for primary cancer.

30

The methods described so-far relate to the determination of the presence or absence of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. By quantitatively measuring the amount of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from an individual, it is possible to predict the progression of B-CLL in an individual.

35

In one embodiment the quantitative measurement is performed during treatment to estimate the efficiency of such treatment.

40

A preferred embodiment of the above diagnostic and prognostic methods is a method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence of SEQ ID No:15 in a biological

sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

A further preferred embodiment of the above diagnostic and prognostic methods is a
5 method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence of SEQ ID No:16 in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for determining the prognosis of the B-CLL.

10 Yet a further preferred embodiment of the above diagnostic and prognostic methods is a method for detecting the presence or absence an expression product, wherein said at least one expression product comprise the nucleotide sequence spanning the junction sequence between Exon-2 (SEQ ID No:15) and Exon-3 (SEQ ID No:16) in a biological sample isolated from an individual for establishing a differential diagnosis of B-CLL or for
15 determining the prognosis of the B-CLL.

The nucleotide sequence spanning the junction between Exon-2 and Exon-3 is the last 20 nucleotides of the 3'-end of SEQ ID No:15 and the first 20 nucleotides of the 5'-end of SEQ ID No:16.

20

The "junction sequence" between two nucleotide sequences, such as two exons, is herein defined as the at least 20 3'-nucleotides of the first exon which is located 5' relative to the second exon and the at least 20 5'-nucleotides of the second exon which is located 3' relative to the first exon.

25

For all diagnostic applications of the present invention, the biological sample may be selected from the group comprising blood, serum, plasma, urine, saliva, lymph node biopsy, bone marrow, spinal liquid, spleen biopsy, and liver biopsy. The cells to be assessed in a sample are preferably leukocytes, mononuclear leukocytes or lymphocytes or
30 B-lymphocytes.

A further embodiment of the present invention also includes a diagnostic kit for *ex vivo* or *in situ* diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) in a individual, the kit comprising a detector molecule capable of detecting the presence or
35 absence of at least one expression product, wherein said at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from the individual.

40 The detector molecule is preferably a nucleotide and even more preferably a nucleotide capable of hybridising to a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 under stringent condition.

B-CLL therapy

With the identification of a new sub-type of B-CLL having a poor prognosis, the present inventors also provide methods for treatment of B-CLL in such patients. This method is based on the finding that transcription products comprising these sequence products are
5 present in B-CLL cells of patients with the poor prognosis. By modifying the activity and/or level of these transcription products, a treatment and/or cure for B-CLL is provided.

Accordingly, in a therapeutic aspect of the present invention there is provided a method of treating a B-CLL sub-type with poor prognosis comprising administering to an individual
10 with a poor prognosis B-CLL diagnosis a compound capable of decreasing or inhibiting the formation of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

15 One such method is based on administering an oligonucleotide capable of inhibiting transcription from SEQ ID No 1 and/or SEQ ID No 5. Said oligonucleotide may comprises at least 8-10 consecutive nucleotides from the sequence of SEQ ID No 1. These sequences constitute the putative promoter sequences controlling the transcription transcription products which comprise a nucleotide sequence selected from the group consisting of SEQ
20 ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. The oligonucleotides bind specifically to the promoter sequences and inhibit transcription of the gene. Such oligonucleotides may comprises nucleotide monomers selected from the group: DNA, RNA, LNA, PNA, methylated DNA, methylated RNA, more preferably PNA or LNA.

25 In a more preferred embodiment the therapeutic methods comprise administering an oligonucleotide capable of binding to a transcriptional product and preventing translation by destroying the transcriptional product. One particularly preferred embodiment of this aspect is RNA interference (RNAi) oligonucleotides.

30 The discovery of the phenomenon RNAi has revealed an entirely new level of gene regulation in eukaryotic cells. It is based on the observation that the presence of long double stranded RNA (dsRNA) in a cell almost completely eliminates the expression of the gene having the same sequence, whereas expression of other unrelated genes are left
35 undisturbed. Although this observation had been know for time in plants as posttranscriptional gene silencing (PTGS) it was not before it was characterised as a general mechanism throughout the animal kingdom that its potentials were fully appreciated. Over the last few years it has been developed as a robust technique to knock down any desirable gene in worms and flies, and quickly a large body of information was
40 gathered about the function of genes in these organisms. Due to the activation of the interferon system by long dsRNA the RNAi method was at that time not applicable in a mammalian system.

A key observation that allowed the harnessing of RNAi as a tool for regulating gene expression in mammals was the observation that chemically synthesised oligo-mer small interfering RNAs (siRNA) effectively suppress gene expression in several human cell lines without inflecting interferon response. This has triggered new promises for siRNA as a
 5 therapeutic drug in humans.

RNAi works by hybridising specifically to the mRNA transcribed by the cell to form a (partly) double stranded RNA molecule. This is recognised as a double stranded molecule by the cell's own nucleases, which degrade them.

10

In order for the technique to work efficiently, the siRNA oligonucleotide comprises a sequence of 5-30 consecutive nucleotides which is the complementary sequence of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18. By
 15 targeting at least one of the nucleotide sequences selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, the transcriptional products characterised by the nucleotide sequences from the group of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 will be eliminated. Example 5 shows that
 20 cells characteristic for the poor prognosis B-CLL sub-type can be eliminated by destroying the herein mentioned transcription products.

By "complementarity" is meant a nucleic acid that can form hydrogen bond(s) with other DNA or RNA sequence by either traditional Watson-Crick or other non-traditional types of
 25 base-paired interactions, e.g. Hoogsteen type.

Preferred siRNA molecules of the present invention are between 5 to 30 nucleotides long, such as 8-30 nucleotides long, such as 8-25 nucleotides, e.g. 8-24 nucleotides, e.g. 8-23 nucleotides, e.g. 8-22 nucleotides, e.g. 8-21 nucleotides, such as 8-20 nucleotides, e.g. 9-
 30 23 nucleotides, e.g. 10-23 nucleotides, such as 11-23 nucleotides, e.g. 12-23 nucleotides such as 13-23 nucleotides, e.g. 14-23 nucleotides, e.g. 15-23 nucleotides, such as 16-23 nucleotides, such as 8 nucleotides, 9 nucleotides, 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23
 35 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, such as 30 nucleotides long.

RNAi oligonucleotides may be administered to the cell, or a vector may be transfected into the cells, said vector comprising a promoter region capable of directing the expression of
 40 at least one RNAi oligonucleotide. Due to the very restricted expression of the AMB-1 gene, it is not important only to target the RNAi oligos or the vectors to B-CLL cells.

One way of targeting to blood cells comprises using a heparin receptor for targeting to blood cells.

Another way of targeting the transcriptional products which comprise at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 is to use
5 an antisense construct comprising a promoter sequence capable of directing the transcription of at least part of the antisense equivalent of SEQ ID No 1 or 2 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 specifically to the poor prognosis B-CLL sub-type.

When desired targeting to B-CLL cells can be performed using the CD19 or CD20 receptor.
10 The CD19 receptor is particularly preferred since it internalises its ligand.

In a further therapeutic embodiment the compound is a gene therapy vector comprising a promoter sequence operably linked to a sequence coding for a protein capable of inhibiting cell division in the cell and/or capable of killing the cell, said promoter sequence being a
15 tissue specific promoter capable of directing expression only in B cells, more preferably only in B-CLL cells. One particularly preferred promoter sequence is the extremely cell specific promoter of SEQ ID No:1 or SEQ ID No:5. When this promoter is used targeting of the suicide vector is not very important, since it will only be active in the cells in which AMB-1 is expressed and these are the cells to be targeted by the suicide gene.

20 Deletion studies will determine the exact length of the promoter sequence counted from the transcription start site. Accordingly, the promoter may comprise at least 100 nucleotides of Seq_ID:1 or Seq_ID:5, such as at least 200 nucleotides, for example at least 300 nucleotides, such as at least 400 nucleotides, for example at least 500
25 nucleotides, such as at least 600 nucleotides, for example at least 700 nucleotides, such as at least 800 nucleotides, for example at least 900 nucleotides, such as at least 1000 nucleotides, for example at least 1100 nucleotides, such as at least 1200 nucleotides, for example at least 1300 nucleotides, such as at least 1400 nucleotides, for example at least 1500 nucleotides, such as at least 1600 nucleotides, for example at least 1700 nucleotides,
30 such as at least 1800 nucleotides, for example at least 1900 nucleotides, such as at least 2000 nucleotides, for example at least 2500 nucleotides, such as at least 3000 nucleotides, for example at least 3500 nucleotides, such as at least 5000 nucleotides, for example at least 10,000 nucleotides.

35 The specificity of expression of mRNAs described by the present invention are striking. The RT-PCR data and the Northern blot data using the dot-blot disclose that the mRNAs of the present invention are expressed either at very low levels in other tissues or only in the B-CLL patients where one can detect it.

40 Thus one embodiment of the present invention is the use of the promotor region for use in gene therapy. The promotor is defined as any sequence within SEQ ID No:1 and SEQ ID No:5 that directs the formation of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18, said expression

product being any of the sequences selected from the group consisting of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11 in B-CLL cells or any other cell or tissue types in which any of the sequences are transcribed.

5

One way that one can use the promotor region in gene therapy is to make a gene therapy construct where the promotor drives the expression of a cell suicide gene such as but not limited to the gene for HSV-1 thymidine kinase, the varicella-zoster, virus thymidine kinase gene, E.Coli cytosine deaminase, the nitroreductase gene or the E.Coli Deo gene (Yazawa et al. & Kirn D. et al.). This would allow for selective expression of the suicide genes in B-CLL cells.

Alternatively, the promotor could be used for a selective expression of genes that could have curative effects when expressed in B-CLL cells, but unwanted effects if expressed ubiquitously.

Also, one embodiment of the present invention relates to the use of the promotor region for use in screening assays where the promotor is linked to a reporter gene and transfected into B-CLL cells in which the reporter gene will be expressed. This approach would allow for easy screening for compounds that would turn off the expression of the reporter gene for example by killing the cell.

A presently preferred embodiment relates a gene therapy vector of the present invention comprising an oligonucleotide capable of inhibiting transcription from SEQ ID No 1 and/or SEQ ID No 5, wherein the promoter is a B-CLL specific promoter, which may or may not be operably linked to a protein selected from the group comprising HSV-1 thymidine kinase, the varicella-zoster, virus thymidine kinase gene, E.Coli cytosine deaminase, the nitroreductase gene or the E.Coli Deo gene.

In one embodiment the compound is a therapeutic antibody directed against a polypeptide having the amino acid sequence of SEQ ID No 3, preferably wherein said antibody is a human or humanised antibody. Another possibility is to identify a modulator of binding of SEQ ID No 3 to its receptor within or outside the cell and to administer this modulator to the cells.

35

4-helical cytokines

A further object of preferred embodiments of the present invention is an isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

40

- i) an amino acid sequence of SEQ ID NO: 3,
- ii) an amino acid sequence having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3

- ii) a fragment of SEQ ID NO:3 having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3.

The protein encoded by SEQ ID No 1 and/or SEQ ID No 5 shares a very small sequence identity with any known protein. However, it has been possible to use 2D and 3D analytical tools to identify the protein as a 4-helical cytokine. The 3D structure of the protein is very similar to 4-helical cytokines and in particular to IL4.

IL4 is a very important cytokine in B-CLL biology. IL4 is not expressed by B-CLL cells, but the IL4 receptor is found on the cells. The IL4 that stimulates B-CLL cells is believed to be produced by T-lymphocytes. The role of IL4 in B-CLL biology is complicated. It has been suggested that IL4 can inhibit B-CLL DNA synthesis and proliferation. Other reports demonstrated that IL4 protects B-CLL cells from apoptosis by upregulating Bcl-2, and IL4 was shown to inhibit apoptosis without stimulating proliferation. Recently, a clinical study in Sweden has confirmed these *in vitro* studies since IL4 administration to B-CLL patients resulted in increased numbers of B-CLL cells in the blood, suggesting that IL4 had a stimulatory or anti-apoptotic effect on the B-CLL cells *in vivo* (Lundin, et al.).

In many systems the effects of IL13 are largely similar to those of IL4, but IL13 is slightly less potent than IL4. It is unclear whether B-CLL cells express IL13, but the cells do express the IL13 receptor. The effects of IL13 in B-CLL are controversial. While Chaouchi et al. suggested that IL13, like IL4 protects B-CLL cells from apoptosis (Chaouchi et al.), studies by Fluckiger et al. suggest that this is not the case (Fluckiger et al.).

The combined finding of 2D and 3D structure similarity to 4-helical cytokines and the importance of IL4 in B-CLL strongly suggests that the novel class of proteins of which the AMB-1 protein is one representative are cytokines.

These polypeptides constitute a novel class of proteins sharing 2D and 3D structure similarities with 4-helical cytokines. In a preferred embodiment, the isolated polypeptide comprises or essentially consists of the amino acid sequence of SEQ ID No. 3 or a fragment thereof. This particular protein at least can be used for diagnosis, for raising antibodies for use in therapy against B-CLL, and for protective or therapeutic immunisation of an individual against B-CLL.

Consequently, the isolated polypeptide preferably has interleukin activity or interleukin like activity, such as having IL3, IL13, GM-CSF, TGF- β , IGF activity, more preferably having IL4 activity or IL4 like activity.

Probably the isolated polypeptides are capable of forming homo- or hetero-oligomer with each other and among themselves. Such oligomers are also within the scope of the present invention. Such oligomers may comprise at least one isolated polypeptides as defined in any the present invention, such as a dimer, a trimer, a tetramer, a pentamer, a

hexamer, an octamer, a decamer, a dodecamer. In biological systems the activity may be attributed only to dimer or higher -mer.

The protein defined by SEQ ID No 3 shares very little sequence identity with known
5 cytokines and interleukines and as a matter of fact very little sequence identity with any
known protein. Consequently the present inventors contemplates that the group comprises
functionally equivalent polypeptide sharing at least 60% sequence identity with SEQ ID No
3, more preferably at least 70% sequence identity, more preferably at least 80 %
sequence identity, such as at least 90 % sequence identity, for example at least 95 %
10 sequence identity, such as at least 97 % sequence identity, for example at least 98 %
sequence identity.

Activity as a cytokine or interleukin can be assessed in a biological assay where the
polypeptide is contacted with a cytokine dependent cell line. Accordingly, polypeptides
15 with cytokine or interleukin like activity can also be identified by similar methods.

One approach to assess cytokine/interleukin activity in a biological assay is to express the
CDS (SEQ ID No:17) reading frame in a baculovirus system (Invitrogen, Carlsbad, USA)
and purify the protein . The recombinant protein can be assayed in cytokine induced
20 proliferation assays as described in general in the eBioscience catalog & Reference Manual
2002 p.260-262 (eBioscience, San Diego, USA). In particular IL4 activity can be
determined using the CTh4S cell line as described by Petersen et al (see Example 8).

The promoter sequence (which forms part of SEQ ID No 1 and/or SEQ ID No 5) and the
25 coding sequences (SEQ ID No:3) can be used in various aspects of gene therapy and
immunotherapy.

Further polynucleotide sequences from other individuals or other species with the same
function can be isolated by one of the following methods, which each form independent
30 aspects of the present invention.

A first method for identifying a nucleotide sequence encoding a 4-helical cytokine
comprises the steps of:

- i) isolating mRNA from a biological sample,
- 35 ii) hybridising the mRNA to a probe comprising at least 10 nucleotides of the coding
sequence of SEQ ID No 1 and/or SEQ ID No 5 under stringent conditions,
- iii) determining the nucleotide sequence of a sequence capable of hybridising under step
ii), and
- iv) determining the presence of an open reading frame in the nucleotide sequence
- 40 determined under step iii).

A second method for identifying a nucleotide sequence encoding a 4-helical cytokine is a
computer assisted method comprising the steps of

- i) performing a sequence similarity search of at least 10 nucleotides of the coding sequence SEQ ID No 1 and/or SEQ ID No 5,
- ii) aligning "hits" to said coding sequence,
- iii) determining the presence of an open reading frame in the "hits".

5

It is highly likely that other similar polypeptides encoding further 4-helical cytokines can be found in other individuals and/or other species of mammals. In particular, individuals of other geographical origin may carry genes which differ from the polynucleotides of the present invention. It is also conceivable that similar sequences can be found in closely and

10 even in distantly related species.

Functional Equivalents

Modification and changes may be made in the structure of the peptides of the present invention and DNA segments which encode them and still obtain a functional molecule that

15 encodes a protein or peptide with desirable characteristics. The following is a discussion based upon changing the amino acids of a protein to create an equivalent, or even an improved, second-generation molecule. The amino acid changes may be achieved by changing the codons of the DNA sequence, according to the genetic code.

20 For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies, binding sites of receptors, or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence

25 substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

30

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporate herein by reference). It is accepted that the relative hydropathic character of the amino

35 acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics (Kyte and Doolittle, 1982), these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine

40 (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, i.e. still obtain a biologically functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those which are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Pat. No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 \pm 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those which are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

Functional equivalents and variants are used interchangeably herein. In one preferred embodiment of the invention there is also provided variants of a 4-helical cytokine, and variants of fragments thereof. When being polypeptides, variants are determined on the basis of their degree of identity or their homology with a predetermined amino acid sequence, said predetermined amino acid sequence being SEQ ID No. 3 or a fragment thereof.

Accordingly, variants preferably have at least 60 % sequence identity, for example at least 65% sequence identity, such as at least 70 % sequence identity, for example at least 75% sequence identity, for example at least 80% sequence identity, such as at least 85 % sequence identity, for example at least 90 % sequence identity, such as at least 91 % sequence identity, for example at least 91% sequence identity, such as at least 92 % sequence identity, for example at least 93 % sequence identity, such as at least 94 % sequence identity, for example at least 95 % sequence identity, such as at least 96 %

sequence identity, for example at least 97% sequence identity, such as at least 98 % sequence identity, for example 99% sequence identity with the predetermined sequence.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % Identity # of identical positions/total # of positions (e.g., overlapping positions) x 100). In one embodiment the two sequences are the same length. The sequence identity is preferably calculated relative to the full length sequence of the molecule of the present invention.

A degree of "sequence identity" of amino acid sequences is a function of the number of identical amino acids at positions shared by the amino acid sequences. A degree of "sequence homology" or "sequence similarity" of amino acid sequences is a function of the number of amino acids, i.e. structurally related, at positions shared by the amino acid sequences. Sequence identity is determined by the alignment algorithm that performs global alignments which has been described by Smith TF and Waterman MS (Smith TF et al.)

A list of the standard qualifiers and the default values for the alignment algorithm is given below:

Standard (Mandatory) qualifiers		Allowed values	Default
[-asequence] (Parameter 1)	Sequence USA	Readable sequence	Required
[-bsequence] (Parameter 2)	Sequence database USA	Readable sequence(s)	Required
-gapopen	The gap open penalty is the score taken away when a gap is created. The best value depends on the choice of comparison matrix. The default value assumes you are using the EBLOSUM62 matrix for protein sequences, and the EDNAFULL matrix for nucleotide sequences.	Number from 1.000 to 100.000	10.0 for any sequence

-gapextend	The gap extension penalty is added to the standard gap penalty for each base or residue in the gap. This is how long gaps are penalized. Usually you will expect a few long gaps rather than many short gaps, so the gap extension penalty should be lower than the gap penalty. An exception is where one or both sequences are single reads with possible sequencing errors in which case you would expect many single base gaps. You can get this result by setting the gap open penalty to zero (or very low) and using the gap extension penalty to control gap scoring.	Number from 0.100 to 10.000	0.5 for any sequence
[-outfile] (Parameter 3)	Output alignment file name	Alignment output file	
Additional (Optional) qualifiers		Allowed values	Default
-datafile	This is the scoring matrix file used when comparing sequences. By default it is the file 'EBLOSUM62' (for proteins) or the file 'EDNAFULL' (for nucleic sequences). These files are found in the 'data' directory of the EMBOSS installation.	Comparison matrix file in EMBOSS data path	EBLOSUM62 for protein EDNAFULL for DNA
Advanced (Unprompted) qualifiers		Allowed values	Default
-[no]brief	Brief identity and similarity	Boolean value Yes/No	Yes

An "unrelated" or "non-homologous" sequence shares less than 40% identity, though preferably less than 25% identity, with one of the 4-helical cytokine sequences of the present invention. The term "substantial identity" means that two peptide sequences, when optimally aligned, such as by the Smith and Waterman algorithm using default gap weights, share at least 80 percent sequence identity, preferably at least 90 percent sequence identity, more preferably at least 95 percent sequence identity or more (e.g., 99 percent sequence identity). Preferably, residue positions which are not identical differ by conservative amino acid substitutions.

Additionally, variants are also determined based on a predetermined number of conservative amino acid substitutions as defined herein below. Conservative amino acid substitution as used herein relates to the substitution of one amino acid (within a predetermined group of amino acids) for another amino acid (within the same group), wherein the amino acids exhibit similar or substantially similar characteristics.

Within the meaning of the term "conservative amino acid substitution" as applied herein, one amino acid may be substituted for another within the groups of amino acids indicated herein below:

- Amino acids having polar side chains (Asp, Glu, Lys, Arg, His, Asn, Gln, Ser, Thr, Tyr, and Cys,)
- Amino acids having non-polar side chains (Gly, Ala, Val, Leu, Ile, Phe, Trp, Pro, and Met)
- 5 • Amino acids having aliphatic side chains (Gly, Ala Val, Leu, Ile)
- Amino acids having cyclic side chains (Phe, Tyr, Trp, His, Pro)
- Amino acids having aromatic side chains (Phe, Tyr, Trp)
- Amino acids having acidic side chains (Asp, Glu)
- Amino acids having basic side chains (Lys, Arg, His)
- 10 • Amino acids having amide side chains (Asn, Gln)
- Amino acids having hydroxy side chains (Ser, Thr)
- Amino acids having sulphur-containing side chains (Cys, Met),
- Neutral, weakly hydrophobic amino acids (Pro, Ala, Gly, Ser, Thr)
- Hydrophilic, acidic amino acids (Gln, Asn, Glu, Asp), and
- 15 • Hydrophobic amino acids (Leu, Ile, Val)

Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

- 20 Accordingly, a variant or a fragment thereof according to the invention may comprise, within the same variant of the sequence or fragments thereof, or among different variants of the sequence or fragments thereof, at least one substitution, such as a plurality of substitutions introduced independently of one another.
- 25 It is clear from the above outline that the same variant or fragment thereof may comprise more than one conservative amino acid substitution from more than one group of conservative amino acids as defined herein above.

- The addition or deletion of at least one amino acid may be an addition or deletion of from
- 30 preferably 2 to 250 amino acids, such as from 10 to 20 amino acids, for example from 20 to 30 amino acids, such as from 40 to 50 amino acids. However, additions or deletions of more than 50 amino acids, such as additions from 50 to 100 amino acids, addition of 100 to 150 amino acids, addition of 150-250 amino acids, are also comprised within the present invention. The deletion and/or the addition may - Independently of one another -
- 35 be a deletion and/or an addition within a sequence and/or at the end of a sequence.

- The polypeptide fragments according to the present invention, including any functional equivalents thereof, may in one embodiment comprise less than 250 amino acid residues, such as less than 240 amino acid residues, for example less than 225 amino acid residues,
- 40 such as less than 200 amino acid residues, for example less than 180 amino acid residues, such as less than 160 amino acid residues, for example less than 150 amino acid residues, such as less than 140 amino acid residues, for example less than 130 amino acid residues, such as less than 120 amino acid residues, for example less than 110 amino acid residues, such as less than 100 amino acid residues, for example less than 90 amino acid residues,

such as less than 85 amino acid residues, for example less than 80 amino acid residues, such as less than 75 amino acid residues, for example less than 70 amino acid residues, such as less than 65 amino acid residues, for example less than 60 amino acid residues, such as less than 55 amino acid residues, for example less than 50 amino acid residues.

5

"Functional equivalency" as used in the present invention is according to one preferred embodiment established by means of reference to the corresponding functionality of a predetermined fragment of the sequence.

- 10 Functional equivalents or variants of a 4-helical cytokine will be understood to exhibit amino acid sequences gradually differing from the preferred predetermined 4-helical cytokine, as the number and scope of insertions, deletions and substitutions including conservative substitutions increases. This difference is measured as a reduction in homology between the preferred predetermined sequence and the fragment or functional
15 equivalent.

- All fragments or functional equivalents of SEQ ID No. 3 are included within the scope of this invention, regardless of the degree of homology that they show to the respective, predetermined 4-helical cytokines disclosed herein. The reason for this is that some
20 regions of the 4-helical cytokines are most likely readily mutable, or capable of being completely deleted, without any significant effect on the binding activity of the resulting fragment.

- A functional variant obtained by substitution may well exhibit some form or degree of
25 native cytokine activity, and yet be less homologous, if residues containing functionally similar amino acid side chains are substituted. Functionally similar in this respect refers to dominant characteristics of the side chains such as hydrophobic, basic, neutral or acidic, or the presence or absence of steric bulk. Accordingly, in one embodiment of the invention, the degree of identity is not a principal measure of a fragment being a variant or functional
30 equivalent of a preferred predetermined fragment according to the present invention.

- One particularly preferred method of determining the degree of functional equivalence is by performing a biological or chemical assay such as the assays described in the appended examples. Preferred functional equivalents of SEQ ID No 3 are those that have a K_D with
35 respect to a predefined receptor which is less than 10 times higher than the K_D of the polypeptide of SEQ ID No 1 with respect to the same receptor, more preferably less than 5 times higher, more preferably less than 2 times higher.

- With respect to functional equivalence this may be defined in a biological assay based on a
40 cytokine dependent or stimulated cell line. Such cell lines are e.g. available from American Type Culture Collection, P.O.Box 1549, Manassas, VA 20108 USA. The following cell lines at least are available for testing cytokines and in particular interleukins:

Accession number	Description	Activity
CRL-1841	TH-2 clone A5E	IL2 dependent, IL4 stimulated
CRL-2003	TF-1	IL3 dependent
CRL-2407	NK-92	IL2 dependent
5 CRL-2408	NK-92MI	IL2 dependent
CRL-2409	NK92CI	IL2 dependent
CRL-9589	AML-193	IL3 stimulated, GM-CSF dep.
CRL-9591	MV-4-11	GM-CSF dependent
TIB-214	CTLL-2	IL2 dependent

10

The following cell lines are available from DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig, GERMANY. As can be seen from the table, some of the cell lines can be used to broadly assess cytokine activity whereas others are only reported to respond to one or a few

15 specific cytokines.

Accession number	Description	Activity
ACC 211	Mouse hybridoma, B9	IL6 dependent
ACC 137	Human acute myeloid leukemia, UT-7	Constitutively cytokine responsive to various cytokines.
ACC 104	Human acute megakaryoblastic leukemia	Respond with proliferation to: GM-CSF, IFN-alpha, IFN- γ , IFN-gamma, IL2, IL3, IL4, IL6, IL15, NG F, SCF, TNF-alpha, TPO
ACC 247	Human acute myeloid leukemia, OCI-AML5	G-CSF, GM-CSF, IL3, FTL3-ligand
ACC 271	Human acute myeloid leukemia, MUTZ-2	IL3, SCF, G-CSF, M-CSF, IFN-gamma
ACC 334	Human erythroleukemia, TF-1	GM-CSF, IFN-gamma, IL3, IL4, IL5, IL6, IL13, LIF, NGF, OSM, SCF, TNF-alpha, and TPO

The TF-1 cell line mentioned above can be used for assaying IL13 function. This cell line is sensitive to various different cytokines but gives a very strong proliferative response when
 20 exposed to IL13. This cell line can in particular be used if there is no response in the IL4 sensitive cell line (CT.h4S). Further cell lines which can be used for distinguishing between IL4 and IL13 activity include cell lines/hybridomas such as B-9-1-3 (Bouteiller et al.) and A201.1 (Andrews et al.).

Pharmaceutical uses of isolated polypeptides

Apart from being used for diagnosis, it is also within the scope of the present invention to use an isolated polypeptide as defined in the invention for a pharmaceutical composition together with a pharmaceutically acceptable carrier. Such pharmaceutical compositions
5 may be used for any of the purposes for which cytokines and in particular interleukin is used at present.

Examples of such uses include the treatment of bone disorders, inflammation, for lowering blood serum cholesterol, allergy, infection, viral infections, hematopoietic disorders,
10 preneoplastic lesions, immune related diseases, autoimmune related diseases, infectious diseases, tuberculosis, cancer, viral diseases, septic shock, reconstitution of the haematopoietic system, induction of the granulocyte system, pain, cardiac dysfunction, CNS disorders, depression, arthritis, psoriasis, dermatitis, colitis, Crohn's disease, diabetes, in an individual in need thereof.

15

It is also within the scope of the present invention to use an isolated polypeptide according to the invention as an adjuvant or as an immune enhancer, for regulating TH2 immune responses, and for suppressing Th1 immune responses.

20 A further use of an isolated polypeptide of the invention is as a growth factor for administration to cell cultures or as a growth factor for veterinary use, e.g. for stimulating the growth of livestock.

Immunotherapy

25 Having identified a transcriptional and/or translational product of SEQ ID No 1 and/or SEQ ID No 5 as an etiological factor in B-CLL it is also within the scope of the present invention to perform an immunisation of a patient in need thereof against B-CLL, wherein the immunisation generates an immune response in the patient which recognises a translational product of SEQ ID No 2, SEQ ID No 4, SEQ ID No 6, SEQ ID No 7, SEQ ID No
30 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11. A preferred immunotherapy is a vaccination against B-CLL by immunising an individual against a translational product of SEQ ID No 1 and/or SEQ ID No 5. In this way the individual builds up antibodies directed against said translational product and any developing B-CLL will be stopped by these antibodies.

35

Immunisation may be performed in various ways, such as by immunising said individual with at least one isolated polypeptide as defined the present invention and optionally adjuvants and carriers or immunising with an expression construct capable of expressing an isolated polypeptide according to the invention in the cells (DNA vaccination).

40

Another method comprises peptide loading of dendritic cells, or ex vivo expansion and activation of T-cells, or inducing a CTL response that targets cells expressing the polypeptide encoded by SEQ ID No 1 and/or SEQ ID No 5.

Antibodies

Antibodies against any of the polypeptides belonging to the novel class of proteins identified by the present inventors can be produced by any known method of immunisation.

5

In one embodiment, the antibodies are produced in a non-human mammal, or in an insect. If antibodies are to be used for therapy in human beings they are preferably subsequently humanised. In one embodiment, the antibody is formulated into a single-chain antibody.

- 10 In another embodiment, in particular for therapeutic purposes, the host organism is a human being and the antibody is subsequently produced recombinantly in a non-human mammal, such as a mouse. The antibody may also be produced as a monoclonal antibody in a hybridoma. One way of producing a monoclonal antibody is described in US 5,681,729 in which a human lymphocyte producing an antibody is generated by the steps, in the

15 order mentioned, comprising

1. transplanting human lymphocytes to a mouse lacking both functional T and B cells so that said human lymphocytes take in said mouse's body;
 2. immunizing said mouse with a desired antigen so as to generate human lymphocytes producing an antibody specific to said antigen;
 - 20 3. administering to said mouse an antiserum to mouse cells;
 4. recovering lymphocyte containing cells from said mouse;
 5. separating human lymphocytes from the recovered cells by centrifugation; and
 6. separating said human lymphocytes producing said antibody.
 7. immortalizing said human lymphocytes
 - 25 8. cloning the obtained immortalized human-derived lymphocytes producing said antibody; and
- recovering a monoclonal antibody specific to said desired antigen from the cloned immortalized human-derived lymphocytes.

- 30 The antibodies of the present invention may be provided as part of a pharmaceutical composition. Such a pharmaceutical composition may be used for treating cancer, preferably for treating leukaemia, more preferably for treating B-CLL leukaemia, more preferably for treating poor prognosis B-CLL leukaemia.

35

Use of antibodies in therapy

- Antibodies directed against epitopes can be used for prevention and/or therapy of for example B-CLL. Antigenic epitopes can be used in vaccines to stimulate an immunological response in a mammal that is directed against cells having the B-CLL-associated epitope
- 40 found in the AMB-1 protein(s) or functional equivalents. Antibodies directed against the antigenic epitopes of the invention can combat or prevent B-CLL.

An antigenic epitope may be administered to the mammal in an amount sufficient to stimulate an immunological response against the antigenic epitope. The antigenic epitope

may be combined in a therapeutic composition and administered in several doses over a period of time that optimizes the immunological response of the mammal. Such an immunological response can be detected and monitored by observing whether antibodies directed against the epitopes of the invention are present in the bloodstream of the
5 mammal.

Such antibodies can be used alone or coupled to, or combined with, therapeutically useful agents. Antibodies can be administered to mammals suffering from any B-CLL that displays the B-CLL-associated epitope. Such administration can provide both therapeutic
10 treatment, and prophylactic or preventative measures. For example, therapeutic methods can be used to determine the spread of a B-CLL and lead to its remission.

Therapeutically useful agents include, for example, leukeran, adrimycin, aminoglutethimide, aminopterin, azathioprine, bleomycin sulfate, busulfan, carboplatin, carminomycin, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine,
15 cytarabine, cytosine arabinoside, cytosine dacarbazine, dactinomycin, daunomycin, daunorubicin, doxorubicin, esperamicins, etoposide, fluorouracil, ifosfamide, interferon- α , lomustine, melphalan, mercaptopurine, methotrexate, mitomycin C, mitotane, mitoxantrone, procarbazine HCl, taxol, taxotere (docetaxel), teniposide, thioguanine, thiotepa, vinblastine sulfate, vincristine sulfate and vinorelbine. Additional agents include those disclosed in Chapter 52, Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner), and the introduction thereto, pp.1202-1263, of Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division). Toxins can be proteins such as, for example, pokeweed anti-viral
20 protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin, or Pseudomonas exotoxin. Toxin moieties can also be high energy-emitting radionuclides such as cobalt-60, I-131, I-125, Y-90 and Re-186, and enzymatically active toxins of bacterial, fungal, plant or animal origin, or fragments thereof.

30 Chemotherapeutic agents can be used to reduce the growth or spread of B-CLL cells and tumors that express the AMB-1 associated epitope of the invention. Animals that can be treated by the chemotherapeutic agents of the invention include humans, non-human primates, cows, horses, pigs, sheep, goats, dogs, cats, rodents and the like. In all embodiments human B-CLL antigens and human individuals are preferred.

35

Species-dependent antibodies can be used in therapeutic methods. Such a species-dependent antibody has constant regions that are substantially non-immunologically reactive with the chosen species. Such species-dependent antibody is particularly useful for therapy because it gives rise to substantially no immunological reactions. The species-
40 dependent antibody can be of any of the various types of antibodies as defined above, but preferably is mammalian, and more preferably is a humanized or human antibody.

Compositions

Therapeutically useful agents can be formulated into a composition with the antibodies of the invention and need not be directly attached to the antibodies of the invention.

However, in some embodiments, therapeutically useful agents are attached to the
5 antibodies of the invention using methods available to one of skill in the art, for example, standard coupling procedures.

Compositions may contain antibodies, antigenic epitopes or trypsin-like protease inhibitors. Such compositions are useful for detecting the AMB-1 protein (for example antigenic
10 epitopes) and for therapeutic methods involving prevention and treatment of B-CLLs associated with the presence of the AMB-1 (for example antigenic epitopes).

The antibodies, (and for example antigenic epitopes and protease inhibitors) can be formulated as pharmaceutical compositions and administered to a mammalian host, such
15 as a human patient in a variety of forms adapted to the chosen route of administration. Routes for administration include, for example, intravenous, intra-arterial, subcutaneous, intramuscular, intraperitoneal and other routes selected by one of skill in the art.

Solutions of the antibodies, (and for example antigenic epitopes and protease inhibitors)
20 can be prepared in water or saline, and optionally mixed with a nontoxic surfactant. Formulations for intravenous or intra-arterial administration may include sterile aqueous solutions that may also contain buffers, liposomes, diluents and other suitable additives.

The pharmaceutical dosage forms suitable for injection or infusion can include sterile
25 aqueous solutions or dispersions comprising the active ingredient that are adapted for administration by encapsulation in liposomes. In all cases, the ultimate dosage form must be sterile, fluid and stable under the conditions of manufacture and storage.

Sterile injectable solutions are prepared by incorporating the antibodies, antigenic epitopes
30 and protease inhibitors in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filter sterilization.

Polynucleotides

In a still further aspect the invention relates to an isolated polynucleotide selected from
35 the group consisting of:

- i) a polynucleotide comprising nucleotides of SEQ ID No 5,
- ii) a polynucleotide encoding a polypeptide having the amino acid sequence of SEQ ID No 3,
- 40 iii) a polynucleotide, the complementary strand of which hybridises, under stringent conditions, with a polynucleotide as defined in any of i) and ii).
- iv) a polynucleotide which is degenerate to the polynucleotide of iii), and
- v) the complementary strand of any such polynucleotide.

The polypeptides encoded by the polynucleotides may furthermore

- a) have at least 60 % sequence identity with the amino acid sequence of SEQ ID No 3 and have interleukin or cytokine activity,
- b) be recognised by an antibody, or a binding fragment thereof, which is capable of recognising an epitope, wherein said epitope is comprised within a polypeptide having the amino acid sequence of SEQ ID No 3; and/or
- c) be competing with a polypeptide having the amino acid sequence as shown in SEQ ID No 3 for binding to at least one predetermined binding partner such as a cytokine receptor.

10

Specific examples of fragments of SEQ ID No 1 include the nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

15 Hybridisation

The entire nucleotide sequence of the coding sequence of SEQ ID No 1 and/or SEQ ID No 5 or portions thereof can be used as a probe capable of specifically hybridising to corresponding sequences. To achieve specific hybridisation under a variety of conditions, such probes include sequences that are unique and are preferably at least about 10 nucleotides in length, and most preferably at least about 20 nucleotides in length. Such probes can be used to amplify corresponding sequences from a chosen organism or individual by the well-known process of polymerase chain reaction (PCR) or other amplification techniques. This technique can be used to isolate additional nucleotide sequences from a desired organism or as a diagnostic assay to determine the presence of the coding sequence in an organism or individual. Examples include hybridisation screening of plated DNA libraries (either plaques or colonies; see e. g. Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, eds., Academic Press).

The terms "stringent conditions" or "stringent hybridisation conditions" include reference to conditions under which a probe will hybridise to its target sequence, to a detectably greater degree than other sequences (e. g., at least twofold over background). Stringent conditions are target sequence dependent and will differ depending on the structure of the polynucleotide. By controlling the stringency of the hybridisation and/or washing conditions, target sequences can be identified which are 100% complementary to a probe (homologous probing).

Alternatively, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing).

Generally, probes for hybridisation of this type are in a range of about 1000 nucleotides in length to about 250 nucleotides in length.

An extensive guide to the hybridisation of nucleic acids is found in Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes,

- Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and Current Protocols in Molecular Biology, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). See also Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual (2nd ed. 5 Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

Specificity is typically the function of post-hybridisation washes, the critical factors being the ionic strength and temperature of the final wash solution.

- 10 Generally, stringent wash temperature conditions are selected to be about 5°C to about 2°C lower than the melting point (T_m) for the specific sequence at a defined ionic strength and pH. The melting point, or denaturation, of DNA occurs over a narrow temperature range and represents the disruption of the double helix into its complementary single strands. The process is described by the temperature of the midpoint of transition, T_m, 15 which is also called the melting temperature.

Formulas are available in the art for the determination of melting temperatures.

- Preferred hybridisation conditions for the nucleotide sequence of the invention include 20 hybridisation at 42°C in 50% (w/v) formamide, 6X SSC, 0.5% (w/v) SDS, 100 mg/ml salmon sperm DNA. Exemplary low stringency washing conditions include hybridization at 42°C in a solution of 2X SSC, 0.5% (w/v) SDS for 30 minutes and repeating. Exemplary moderate stringency conditions include a wash in 2X SSC, 0.5% (w/v) SDS at 50°C for 30 minutes and repeating.

- 25 Exemplary high stringency conditions include a wash in 2X SSC, 0.5% (w/v) SDS, at 65°C for 30 minutes and repeating. Sequences that correspond to the AMB-1 gene or fractions thereof according to the present invention may be obtained using all the above conditions. For purposes of defining the invention, the high stringency conditions are used.

30

Mutations

- Finally, the invention provides a method for determining an increased or decreased predisposition for B-CLL comprising determining in a biological sample from an individual a germline alteration in a target nucleic acid sequence comprising 150,000 nucleotides, said 35 target nucleic acid sequence comprising at least 10 nucleotides of SEQ ID No 1 and/or SEQ ID No:5. This aspect is based on the finding of the importance of the expression product of SEQ ID No 1 and/ SEQ ID No:5, and the complete absence of any detectable expression product of SEQ ID No 1 and/or SEQ ID No:5 in healthy tissue and in patients with good prognosis B-CLL. It is highly likely that the difference is caused by a germline alteration. A 40 germline alteration can be targeted by gene therapy methods and by the methods provided in the present invention.

Preferably, said predisposition is a predisposition for poor prognosis of B-CLL.

Examples

Example 1: cDNA cloning

By Differential Display (Pardee et al., 1992, Jørgensen et al., 1999) part of a gene (hereafter referred to as AMB-1) was found that is expressed in unmutated B-CLL patients with poor prognosis. This gene is not found in the mutated B-CLL patients. When AMB-1 was sequenced and aligned to known sequences in GenBank, perfect homology was found to 225 base pairs (bp) of human genomic DNA from chromosome 12.

RNA was prepared using the RNeasy kit from Qiagen, as described by the manufacturer (Qiagen, Hilden, Germany). RNA was prepared from patients with B-CLL without hypermutation who, by PCR analysis, using primers FDP5 (CCTTTATGTGTGTGACAAGTG) and F10 (ATCCAGCCAGGATGAAATAGAA), showed a high level of the resulting PCR fragment. Poly-A⁺ RNA was isolated from total RNA by the "MicroPoly(A)Purist" kit from Ambion, as described by the manufacturer (Ambion, Inc., Texas, USA). Cloning-ready cDNA was prepared from 8 µg poly-A⁺ RNA using the "ZAP Express® XR Library Construction Kit" from Stratagene as described by the manufacturer (Stratagene, San Diego, USA). The cDNA was size fractionated and two size fractions (fraction-1: > 2500 bp and fraction-2 300-2500 bp) were independently ligated to pre-digested lambda Zap vectors and packed into phage particles as described by the manufacturer (Stratagene, San Diego, USA). The titer was determined for each library and 200,000 pfu of from the fraction-1 library were plated onto two 22x22 cm screening plates (100,000 pfu on each plate) and 750,000 pfu of the fraction-2 library were plated on five 22x22 cm screening plates (150,000 pfu on each) as described by Stratagene, San Diego, USA. The plates were incubated at 37°C for 18 hours and the plaques transferred to replica nylon filters (Amersham) and denatured and renatured to allow hybridisation. All procedures were made as described by the manufacturers (Stratagene, San Diego, USA & Amersham Biosciences, Buckinghamshire, UK).

The filters were screened by independent hybridisations with alpha[³²P]-dATP-labelled DNA fragments; alpha[³²P]-dATP was purchased from Amersham Biosciences, Buckinghamshire, UK. Between successive hybridisations, the old probe was removed by incubation for 20 min in 2l 90-100°C water containing 0.1% SDS. The DNA fragments used as probes were (all positions relate to sequence ID # X): 1) pos. 48978-49250; 2) pos. 50011-51591; 3) pos. 51461-52182; 4) pos. 51901-52589; 5) pos. 53121-56521; 6) pos. 58163-59408. All hybridisations and washes were made according to the instructions from Stratagene, San Diego, USA and Amersham Biosciences, Buckinghamshire, UK; Washing was done at a high stringency (0.1 x SSC at 65°C for 20 min).

A total of 38 plaques that showed a positive response from one or more of the screenings were excised from the screening plates and grown as plasmids as described (Stratagene, San Diego, USA).

A total of 8 cDNAs were identified by cDNA cloning or by a combination of cDNA cloning, PCR analysis and RACE (rapid amplification fo cDNA ends-polymerase chain reaction) using the the SMART RACE cDNA amplification kit (Clontech, Palo Alto, CA) according to the manufacturer's instructions.

5

Example 2: Bioinformatic analysis of AMB-1.

The DNA and protein sequence data bases (GenBank and EBI) have been searched for sequences with similarity to AMB1. There is no significant DNA sequence similarity to any known gene. In particular, the coding region of the AMB1 mRNA (SEQ ID no 3) is not
10 present in any known EST. The only significant match to the complete mRNA sequences and the DNA sequence of the putative coding region were BAC clones derived from the region on human chromosome 12 where the gene is located. The "AMB-1 gene" had not been annotated as a gene on the chromosome. Searches with the peptide sequence in the sprtnr data base of peptide sequences (includes Sprot and nrtrembl) showed a low
15 similarity to putative intron maturases from cloroplasts and to bovine IL4. The percentage similarity to both maturases and bovine IL4 was low (25.6% and 30.3%, respectively) and the similarity to maturases only included a match to 75 amino acids of the much larger maturases. In contrast, the match to bovine IL4 extended over the full peptide sequence. IL4, and other 4-helical cytokines, include a leader peptide sequence (signal peptide)
20 allowing the proteins to be secreted. The AMB1 peptide sequence includes a N-terminal peptide sequence with similarity to signal peptide sequences, however, it is not a typical sequence.

A 3D search has been performed, where a peptide sequence is searched for similarity to
25 known protein or peptide 3D-structures. The two best matches were the thioredoxin fold and the human 4-helical cytokine IL4 (Fig. 5). The two matches had almost similar probability scores (2.88 and 3.05, respectively). Searches with 4-helical cytokine peptide sequences (IL4, IL3, IL13 and GM-CSF) revealed that all could be folded into both a 4-helical cytokine structure and the thioredoxin fold. Alignment based on the structural
30 similarity between IL4 (d1iara) and AMB-1 is shown in Fig. 5. Thus, the AMB1 peptide sequence share this property with 4-helical cytokines. The structural similarity is not perfect (Fig. 6) and there are no obvious glycosylation sites in the AMB1 sequence, however, the similarity is significant. Alignment of the AMB1 peptide sequence with the sequences of IL4, IL3, IL13 and GM-CSF, based on their structures, showed very little
35 sequence conservation but a high degree of structural conservation (Fig.7). Based on this alignment, AMB1 has similarities to all the 4-helical cytokines, and the length of AMB1 and the position of gaps in the alignment could suggest a higher similarity to e.g. IL13, but searches at 3D-PSSM only identified a significant similarity to the structure of IL4, not IL13, IL3 or GM-CSF. However, the search algorithms are not perfect and may therefore
40 not detect a possible low structural similarity.

Example 3: Differential expression of AMB-1

Patient material

Blood samples were collected from newly diagnosed untreated patients with B-CLL.

- 5 Mononuclear cells were isolated by Lymphoprep separation (Nycomed Pharma, Oslo, Norway), and the percentage of CD5+CD20+ B-CLL cells in the mononuclear fraction was >90% in all samples as determined by flow cytometric analysis.

Isolation of RNA and conversion to cDNA.

- 10 Material for RNA production was isolated mononuclear cells from B-CLL patients or mononuclear cells from lymphoprep separated buffy coats from normal donors. Total RNA was isolated from 5×10^7 or more cells using the QIAamp RNA Blood Mini kit (Qiagen, Valencia, CA) with DNase treatment. RNA (1 μ g) was converted to cDNA by incubation with a mixture of random-primers (1 μ g) and T24-primer (1 μ g) for 5 minutes at 70°C. After
- 15 cooling on ice, the reaction mixture was added to a final volume of 25 μ l containing 30U of AMV Reverse Transcriptase HC (Promega, Madison, WI, USA), 1x First Strand Buffer (50mM Tris-HCl, pH 8.3, 50mM KCl, 10mM MgCl₂, 10mM DTT, 0.5mM spermidine), 2.5mM of each dNTP and 60U rRNasin ribonuclease inhibitor (Promega, Madison, WI, USA). The reaction was performed for 60 minutes at 37°C.

20

Determination of somatic hypermutation status

Two μ l of cDNA was amplified using a GeneAmp PCR System 2700 (Applied Biosystems, Warrington, UK) with a 40 pmol specific upstream primer corresponding to 1 of the 6 human VH family leader sequences (VH1: 5'-CCATGGACTGGACCTGGAGG-3', VH2: 5'-

- 25 ATGGACATACTTTGTTCCAGC-3', VH3: 5'-CCATGGAGTTTGGGCTGAGC-3', VH4: 5'-ATGAAACACCTGTGGTTCTT-3', VH5: 5'-ATGGGGTCAACCGCGATCCT-3', VH6: 5'-ATGTCTGTCTCCTTCCTCAT-3') and a 40 pmol downstream primer (C μ : 5'-GAGGCTCAGCGGAAGACCTT-3' or C γ : 5'-GGGGAAGACCGATGGGCCCT-3') corresponding to a consensus sequence of the constant region of IgM or IgG respectively. The Reverse
- 30 Transcription (RT)-PCR reaction contained 1xPCR buffer (10mM Tris-HCl, pH 9.0, 50mM KCl, 0.1% Triton X-100), 2.5mM MgCl₂, 0.2mM of each dNTP and 1.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a final volume of 100 μ l. The RT-PCR was performed under the following conditions: 1 cycle of 94°C for 5 minutes, 30 cycles of denaturation at 94°C for 30 secs, annealing at 62°C for 30 sec. and extension at 72°C for
- 35 30 sec, and a final extension at 72°C for 7 minutes. The RT-PCR products were analysed on 2% agarose gels and sequenced in an HBI Prism 310 Genetic Analyzer (Perkin Elmer, Foster City, CA, USA) using the BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Warrington, UK) following the manufacturer's instructions.

- 40 Sequences obtained from each sample were compared to germ line sequences in the V base sequence directory (I.M. Tomlinson, MRC Center for Protein Engineering, Cambridge, UK) using BLAST, and the closest germ line sequence was assigned. A gene sequence was

considered to be mutated if it had equal or more than 2% sequence alterations when compared to the closest published germ line sequence.

RT-PCR that amplifies the Exon 2-Exon 3 junction

- 5 To evaluate the mRNA expression pattern of AMB1 in unmutated and mutated B-CLL patients RT-PCR was performed. Exon-overlapping oligonucleotide primers were: 5'-ATCCAGCCAGGATGAAATAGAA-3' and 5'-CACTTGTCACACACATAAAGG-3'. The RT-PCR was performed in a GeneAmp PCR System 2700 thermal cycler with an initial denaturation at 94°C for 2 minutes, 40 cycles of 96°C for 25 sec., 62°C for 25 sec. and 72°C for 90 secs, and a final extension at 72°C for 5 minutes. The reactions contained 2µl cDNA, 1x DDRT-PCR buffer (10mM Tris-HCl, pH 8.3, 50mM KCl, 1.8mM MgCl₂, 0.1% Triton X-100, 0.005% gelatine), 0.25mM of each dNTP, 30 pmol of each primer and 0.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a 30µl final volume. RT-PCR products were analyzed by gelelectrophoresis on 2% agarose gels and visualized with a Gene Genius Bio Imaging System (Syngene, Frederick, MD) after staining with ethidium bromide. An actin control RT-PCR was performed using the primers: 5'-TGACGGGGTCACCCACACTGTGCCCATCTA-3' and 5'-CTAGAAGCATTTGCGGTGGACGATGGAGGG-3'.

RT-PCR that amplifies the Exon 1-Exon 3 junction:

- 20 To evaluate the mRNA expression pattern of AMB1 in unmutated and mutated B-CLL patients RT-PCR was performed. Exon-overlapping oligonucleotide primers were: 5'-AGACGGCTCTCACC AATAAG-3' and 5'-CACTTGTCACACACATAAAGG-3'. The RT-PCR was performed in a GeneAmp PCR System 2700 thermal cycler with an initial denaturation at 94°C for 2 minutes, 40 cycles of 96°C for 25 sec., 62°C for 25 sec. and 72°C for 90 secs, and a final extension at 72°C for 5 minutes. The reactions contained 2µl cDNA, 1x DDRT-PCR buffer (10mM Tris-HCl, pH 8.3, 50mM KCl, 1.8mM MgCl₂, 0.1% Triton X-100, 0.005% gelatine), 0.25mM of each dNTP, 30 pmol of each primer and 0.5U Taq DNA Polymerase (Promega, Madison, WI, USA) in a 30µl final volume. RT-PCR products were analyzed by gel electrophoresis on 2% agarose gels and visualized with a Gene Genius Bio Imaging System (Syngene, Frederick, MD) after staining with ethidium bromide (Continental Lab Products, San Diego, USA).

Statistical analysis

- Statistical significance of the correlation between somatic hypermutation status and AMB1 expression was analyzed using Fisher's exact test.

- Northern blotting.** RNA from spleen, bone marrow and colon was purchased from Clontech. The AMB1 probe was an 896 base pair fragment (57661-56766) obtained by RT-PCR as described above with the primers 5'-TCACCTGGGAGCTCAGAGGA-3' and 5'-GTGATCCTGGGAGAATCTCT-3'. For Northern blotting, 5 µg of RNA was run on a 1% agarose-gel with 6% formaldehyde dissolved in 1 x MOPS (20 mM 3-(N-morpholino)-propane-sulfonic acid, 5 mM sodium acetate, 1 mM EDTA, pH 7.0) for size separation. The presence of equal amounts of RNA in each lane was ensured by ethidium bromide staining. The RNA was transferred to a Hybond-N membrane (Amersham, Little Chalfont, UK) by

capillary blotting and fixed by UV-irradiation. The filters were pre-hybridized for 1-2 hours at 42°C in 6 ml ULTRAhyb (Ambion, Austin, TX, USA) preheated to 68°C and hybridized overnight at 42°C after addition of further 4 ml containing the ³²P-labeled probe and sheared salmon sperm DNA (10 µg/ml). The membranes were washed for 2 x 15 min. at 42°C in 2 x SSC, 0.1% SDS followed by 1 x 15 min. in 0.2 x SSC, 0.1 % SDS and 2 x 15 min. in 0.1 x SSC, 0.1 % SDS at 42°C. The blot was developed and quantified by a phosphoimager. The sizes of the mRNAs were determined by reference to 18S and 28S ribosomal RNA, which were visualized by ethidium bromide staining. The AMB1 probe used for hybridization was radiolabeled with [α -³²P] dCTP using the Random Primers DNA Labeling System (Gibco BRL).

Dot blot of multiple tissue expression (MTE) array. An MTE array (Clontech, Palo Alto, CA, USA) was hybridised to AMB1 at 65°C in ExpressHyb (Clontech) supplemented with sheared salmon sperm DNA (7.5 µg/ml) and human C₀t-1 DNA (1.5 µg/ml) according to the manufacturers recommendations. The tissue types represented on the MTE array are shown in Figure 11. Following hybridisation the filter was washed 5 x 20 min. at 65°C in 2 x SSC (1 x SSC = 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 1% SDS and 2 x 20 min at 65°C in 0.1 x SSC, 0.5% SDS. The blot was developed and quantified by a phosphoimager (Fuji Imager Analyzer BAS-2500, Image Reader ver. 1.4E, Image Gauge ver. 3.01 software, Fuji, Stockholm, Sweden). The membranes were stripped by boiling in 0.5% SDS for 10 min. before rehybridization. The probe used for hybridization were radiolabeled with [α -³²P] dCTP using the Random Primers DNA Labeling System (Gibco BRL, Rockville, MD, USA).

25 Results

Based on the known sequence of the AMB-1 cDNA RT-PCRs with primers that extend across the Exon 2-Exon 3 junction and the Exon 1- Exon 3 junction were set up. As shown in Figure 10, where the Exon 2- Exon 3 junction has been amplified, AMB-1 is expressed in the unmutated patients (UPN1-8) while no expression of AMB-1 is seen in mutated patients (UPN9-16).

Northern blot analysis was performed to determine the size of AMB-1's mRNA transcript. As shown in Figure 3 the probe identifies transcripts in the samples from the three patients without somatic hypermutation (UPN1, UPN4 and UPN7). However, the probe does not recognise any transcripts from the patients with somatic hypermutation (UPN9, UPN10, UPN13, UPN21) or the various cell lines and tissue samples. Similar results were obtained when cell lines and tissue samples were investigated by RT-PCR (results not shown). Dot blot analysis on a purchased filter with 96 different RNA samples (Figure 8) only revealed specific binding to the total DNA control dot, but not to any specific tissue (results not shown). Thus AMB-1 is only expressed in B-CLL cells without hypermutation or AMB-1 is expressed at extremely low levels in other tissues.

We next tested the predictive value, in terms of IgV_H mutational status, of expression of AMB-1 in 29 consecutive newly diagnosed patients. At present 13 somatically unmutated and 16 somatically mutated patients have been included in our prospective patient database. The sensitivity and specificity for expression of AMB-1 in predicting mutational status is well above 90% ($p < 0.0001$), which is at the level obtained by sequencing.

Example 4. Investigation of the prognostic significance of AMB-1 in terms of patient survival

- 10 To obtain information about the ability of AMB-1 to predict survival or time to progression for B-CLL patients survival curves were made. For each patient the following data were collected: clinical stage at time of diagnosis (Rai and Binet staging), date of diagnosis, date of first time of treatment and last follow up date. Time to treatment (progression free survival) and survival time were calculated based on these dates using the Microsoft Excel software. Survival times and progression free survival times were plotted by the Kaplan-Meier Method and compared using the log-rank test (Prism 3.0 Graph Pad software).

34 newly diagnosed, untreated B-CLL patients were investigated as described above. As shown in Figure 2a-d AMB-1 expression is a good predictor of B-CLL patient survival and B-CLL patient time to progression.

Example 5: siRNA assays.

Materials and Methods:

- Frozen cells from unmutated or mutated B-CLL patients purified and characterized for mutational status as previously described were thawed and Ficoll separated to obtain the live cells. Cells were counted and resuspended at 2×10^6 cells/ml in RPMI 1640 medium with glutamax-1 (RPMI, GIBCO, Paisley, UK). 1×10^6 cells (500 μ l of cell suspension) were incubated on ice for 10 minutes with the appropriate siRNA or combination of siRNAs at a final concentration of 100nM each. The suspension was electroporated using a 4mm cuvette (Molecular BioProducts, San Diego, CA, USA) in a BioRad Gene Pulser. Following electroporation the cells were incubated on ice for 10 minutes, washed once in RPMI with 20% FCS and 100units/ml Penicillin and 100 μ g/ml Streptomycin (GIBCO, Paisley, UK), resuspended at a concentration of 4×10^6 cells/ml in the same medium and incubated for 48 hours. For each patient the optimal voltage allowing for maximum siRNA uptake was determined by making a voltage curve (300-450V, 960 μ F) using a FITC-labeled unspecific probe (Xeragon, Qiagen, Hilden, Germany) at a final concentration of 200nM and the uptake was followed by flow cytometry.

The siRNAs were produced using the SilencerTM siRNA Construction Kit, (Ambion, Texas, USA) according to the manufacturers instructions. siRNAs had the following target sequences:

siRNA name:	mRNA target sequence:	
-------------	-----------------------	--

Exon 1 A	5' - AAUAAGGGCAGGCAUCAUCCA -3'	SEQ ID No:19
Exon 1 B	5' - AAUUACACUGCCAGGUUCCU -3'	SEQ ID No:20
Exon 2 A	5' - AAUUCAUUCACAAUGAUUGCU -3	SEQ ID No:21
Exon 2 B	5' - AAUUUCUCUUGGGUAAUUCAG-3'	SEQ ID No:22
Exon 3 (DDend) A	5' - AAAAUCAGAAUCUGCGCAGCA -3'	SEQ ID No:23
Exon 3 (DDend) B	5' - AAUGAUGAUGGGAAGAAGGAA -3'	SEQ ID No:24
CDS A	5' - AAACUUAGUAAUUGAGUGUGA 3'	SEQ ID No:25
CDS B	5' - AAUAUGUCACUUUCAUAAAGC -3'	SEQ ID No:26
Transcript on -strand A	5' - AAUGAUGAUGGGAAGAAGGAA -3'	
Transcript on -strand B	5' - AAACUAUGAGAUUUCAGAAGG -3'	

Table 1: Overview of the siRNAs used in electroporation experiments

After 48 hours of incubation the various B-CLL samples were counted and live and dead cells were distinguished by nigrosin exclusion (0.1% in Nigrosin in PBS from Fluka, Buchs, Switzerland).

Results:

Following 48 hours of incubation with control GFP siRNA or siRNAs against various regions of the cDNAs (see table 1), the viable cells and dead cells were counted based on the ability of the cells to exclude nigrosin. As shown in table 2, in 3 out of 4 patients introduction of siRNAs against Exon 1, Exon 2 and Exon 3 resulted in an increase of dead cells as compared to the controls (no siRNA or GFP siRNA).

15

Treatment:	UPN 67	UPN 66	UPN 62	UPN 73
No siRNA	33.0%	19.7%	30.9%	20.2%
Control siRNA	28.4%	26.7%	24.0%	16.0%
CDS siRNA	23.7%	25.5%	21.0%	13.3%
Ex1-3 siRNA	47.3%	54.1%	17.2%	29.1%

Table 2: Percent dead cells of total cells following electroporation and incubation with siRNAs.

20 Example 6: Identification of possible cytogenetic aberrations near or within the region encoding AMB-1 on chromosome 12.

Rationale: The limited expression profile of AMB-1 suggests that it may be a result of a genetic aberration (e.g. deletion, translocation or alternative splicing) or that the promotor region controlling the expression of AMB-1 is uniquely activated in unmutated B-CLL.

25 Another gene is situated about 200.000 bases upstream of the AMB-1 gene (SEQ ID No 1) on chromosome 12 and the inventors we have determined that this gene is expressed at equal levels in unmutated and mutated patients.

Methods: Using primers, initially spaced about 20.000 bp apart; this region on chromosome 12 is characterised in unmutated B-CLL patients. If genetic aberrations within the region are detected by PCR analysis of chromosomal DNA, detailed molecular genetic studies using FISH, microsatellite analysis and Southern blotting will be employed. The whole region from unmutated patients is sequenced.

Example 7: Polyclonal antibodies

10

Production of polyclonal antibodies:

Synthetic peptides CDLETNSEINKLIYLF SQNNRIRF and CQVSKKHIIYSTFLSKNF were synthesized and conjugated to KLH (K.J.Ross-Petersen Aps, Holte, Denmark). Polyclonal antibodies were produced by immunization of rabbits with these conjugated peptides by DAKO (DAKO Cytomation A/S, Glostrup, Denmark).

Testing of polyclonal antibodies:

At the present time we have produced polyclonal antibodies from three rabbits that have been immunized with peptides representing predicted immunogenic regions of the protein that can be predicted from the CDS sequence (SEQ ID No:17) (cDNA 4). The antibodies are tested in various ways. The proposed reading frame of CDS (SEQ ID No:17) is expressed in 293 cells and the binding of antibodies to lanes on a western blot with non-transfected 293 cells versus transfected 293 cells are compared. The size of the band in the lanes with transfected 293 cells is compared to the size of western blot bands in lanes with proteins from B-CLL patients. Specificity of the bands is secured by peptide blocking experiments.

Additionally the polyclonal antibodies are tested in B-CLL Immunoprecipitation experiments where the antibodies are used to immunoprecipitate the protein produced from the CDS sequence and the immunoprecipitates are analyzed by western blotting.

Example 8: Assay for the biological activity of 4-helix cytokines.

The assay is based on the use of a cytokine dependent or stimulated cell line, for example an IL4 dependent cell line ("Optimisation of the CT.h4S bioassay for detection of human
5 interleukin-4 secreted by mononuclear cells stimulated by phytohaemagglutinin or by human leukocyte antigen mismatched mixed lymphocyte culture", Petersen, S.L., Russell, C.A., Bendtzen, K. & Vindeløv, L.L., Immunology Letters 84 (2002) 29-39). Other examples of cytokine dependent cell lines include IL13 dependent cell lines. A list of commercially available cytokine dependent cell lines is disclosed in the general part of the
10 description. These can all be used for assessing cytokine activity. The most preferred cell lines are those that are IL4 dependent.

The assay can be performed in two ways. The first assay comprises providing recombinantly produced AMB1 protein or a functional equivalent thereof and determine the
15 proliferation rate of the cell line. The proliferation rate (either rate of proliferation or \pm proliferation) can be compared to the proliferation rate of the cell line exposed to IL4 or another known 4-helical cytokine or interleukin.

If a positive result is obtained with a polypeptide an assay will be performed on the same
20 cell line with the IL4 receptor blocked. This will check whether the stimulus goes through IL4R.

The second assay is based on transfection of a gene encoding a 4-helix cytokine according to the invention into cytokine dependent cells and observe proliferation or non-proliferation
25 during transient expression.

Example 9: Cytokine receptor binding assays

The following is a description of the layout of a cytokine receptor binding assay used to determine the cytokine activity of the 4-helix cytokines according to the present invention.
30

The assays can be performed with any cytokine receptor. Preferred receptors include but is not limited to the receptors for IL4 IL13, IL3, and GM-CSF.

The ability of recombinant cytokine receptor to bind to 4-helical cytokine is assessed in a
35 competitive binding ELISA assay as follows. Purified recombinant cytokine receptor (IL4, IL13, IL3 or GM-CSF receptors) (20 μ g/ml in PBS) is bound to a Costar EIA/RIA 96 well microtiter dish (Costar Corp, Cambridge Mass., USA) in 50 μ L overnight at room temperature. The wells are washed three times with 200 μ L of PBS and the unbound sites blocked by the addition of 1% BSA in PBS (200 μ L/well) for 1 hour at room temperature.
40 The wells are washed as above. Biotinylated AMB-1 (1 μ g/ml serially diluted in twofold steps to 15.6 ng/mL; 50 μ L) is added to each well and incubated for 2.5 hours at room temperature. The wells are washed as above. The bound biotinylated AMB-1 is detected by the addition of 50 μ L/well of a 1:2000 dilution of streptavidin-HRP (Pierce Chemical Co.,

Rockford, Ill.) for 30 minutes at room temperature. The wells are washed as above and 50 μ L of ABTS (Zymed, Calif.) added and the developing blue color monitored at 405 nm after 30 min. The ability of unlabelled 4-helical cytokine to compete with biotinylated AMB-1, respectively, is assessed by mixing varying amounts of the competing protein with a
 5 quantity of biotinylated AMB-1 shown to be non-saturating (i.e., 70 ng/mL; 1.5 nM) and performing the binding assays as described above. A reduction in the signal (Abs 405 nm) expected for biotinylated 4-helical cytokine indicates a competition for binding to immobilised cytokine receptor.

- 10 The above identified assays can be used to identify 4-helical cytokines with similar binding affinities as AMB-1 (SEQ ID No. 3). In the competitive binding assays biotinylated IL4, IL13, IL3, or GM-CSF can be used to identify 4-helical cytokines which can compete with these cytokines.

15 References

1. U.S. Pat. No. 4,554,101
2. U.S. Pat. No. 5,681,729
3. Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner
4. Andrews, R., L. Rosa, M. Daines, and G. Khurana Hershey. 2001. Reconstitution of a
 20 functional human type II IL4/IL13 receptor in mouse B cells: demonstration of species specificity. *J Immunol* 166, no. 3:1716
5. Ausubel, et al., Current Protocols in Molecular Biology, Chapter 2, Eds., Greene Publishing and Wiley-Interscience, New York (1995).
6. Bouteiller, C.L., R. Astruc, A. Minty, P. Ferrara, and J.H. Lupker. 1995. Isolation of an
 25 IL13-dependent subclone of the B9 cell line useful for the estimation of human IL13 bioactivity. *J Immunol Methods* 181, no. 1:29
7. Chaouchi, N., C. Wallon, C. Goujard, G. Tertian, A. Rudent, D. Caput, P. Ferrera, A. Minty, A. Vazquez, and J.F. Delfraissy. 1996. Interleukin-13 inhibits interleukin-2-induced proliferation and protects chronic lymphocytic leukemia B cells from in vitro
 30 apoptosis. *Blood* 87, no. 3:1022
8. Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, S.M. Lichtman, P. Schulman, V.P. Vinciguerra, K.R. Rai, M. Ferrarini, and N. Chiorazzi. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 94, no. 6:1840
- 35 9. Dancescu, M., M. Rubio-Trujillo, G. Biron, D. Bron, G. Delespesse, and M. Sarfati. 1992. Interleukin 4 protects chronic lymphocytic leukemic B cells from death by apoptosis and upregulates Bcl-2 expression. *J Exp Med* 176, no. 5:1319
10. Dohner, H., S. Stilgenbauer, A. Benner, E. Leupolt, A. Krober, L. Bullinger, K. Dohner, M. Bentz, and P. Lichter. 2000. Genomic aberrations and survival in chronic
 40 lymphocytic leukemia. *N Engl J Med* 343, no. 26:1910
11. Fluckiger, A.C., F. Briere, G. Zurawski, J.M. Bridon, and J. Banchereau. 1994. IL13 has only a subset of IL4-like activities on B chronic lymphocytic leukaemia cells. *Immunology* 83, no. 3:397

12. Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division)
13. Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 94, no. 6:1848
14. Innis et al. (1990) PCR Protocols, A Guide to Methods and Applications, eds., Academic Press
15. Jørgensen M, Bévort M, Kledal TS, Hansen BV, Dalgaard M, and Leffers H (1999) Differential Display Competitive PCR: An Optimal Tool for Assaying Gene Expression. *Electrophoresis* 20:230-240
16. Kirn D. et al., "The emerging fields of sulcide gene therapy and virotherapy", *Trends Mol. Med.* 2002, 8:568-573
17. Kröber, A., T. Seiler, A. Benner, L. Bullinger, E. Bruckle, P. Lichter, H. Dohner, and S. Stilgenbauer. 2002. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* 100, no. 4:1410
18. Kyte J and Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, no. 1:105
19. Liang P, and Pardee AB (1992) Differential Display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-71
20. Lin, K., P.D. Sherrington, M. Dennis, Z. Matrai, J.C. Cawley, and A.R. Pettitt. 2002. Relationship between p53 dysfunction, CD38 expression, and IgV(H) mutation in chronic lymphocytic leukemia. *Blood* 100, no. 4:1404
21. Lundin, J., E. Kimby, L. Bergmann, T. Karakas, H. Mellstedt, and A. Osterborg. 2001. Interleukin 4 therapy for patients with chronic lymphocytic leukaemia: a phase I/II study. *Br J Haematol* 112, no. 1:155
22. Luo, H.Y., M. Rubio, G. Biron, G. Delespesse, and M. Sarfati. 1991. Antiproliferative effect of interleukin-4 in B chronic lymphocytic leukemia. *J Immunother* 10, no. 6:418
23. Merrifield, J. *Am. Chem. Soc.* 85:2149-2146, 1963)
24. Oscier, D.G., A.C. Gardiner, S.J. Mould, S. Glide, Z.A. Davis, R.E. Ibbotson, M.M. Corcoran, R.M. Chapman, P.W. Thomas, J.A. Copplestone, J.A. Orchard, and T.J. Hamblin. 2002, Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. *Blood* 100, no. 4:1177
25. Panayiotidis, P., K. Ganeshaguru, S.A. Jabbar, and A.V. Hoffbrand. 1993. Interleukin-4 inhibits apoptotic cell death and loss of the bcl-2 protein in B-chronic lymphocytic leukaemia cells in vitro. *Br J Haematol* 85, no. 3:439
26. Petersen, S.L., Russell, C.A., Bendtzen, K. & Vindeløv, L.L., Optimisation of the CT.h4S bioassay for detection of human interleukin-4 secreted by mononuclear cells stimulated by phytohaemagglutinin or by human leukocyte antigen mismatched mixed lymphocyte culture", *Immunology Letters* 84 (2002) 29-39
27. Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.).

28. Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, Part I, Chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993)
29. Yazawa et al., Current Progress In suicide gene therapy for cancer, World J. Surg.
5 2002, 26(7):783-789.

Claims

1. A method for establishing a diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) in a individual comprising detecting the presence or absence of at least
5 one expression product, wherein said at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from the individual.
- 10 2. A method for establishing the prognosis of B-CLL in a individual comprising detecting the presence or absence of at least one expression product, wherein said at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from the individual.
- 15 3. A method for determining whether an individual has a B-CLL sub-type with poor prognosis, the method comprising determining the level of an expression product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18
20 of said individual, and indicating the individual as having a B-CLL sub-type with poor prognosis if the level of the expression product is at or beyond a discriminating value and indicating the individual as not having a B-CLL sub-type with poor prognosis if the level of the expression product is not at or beyond the discriminating value, the discriminating value being a value which has been determined by measuring the level of the expression
25 product which comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in both a healthy control population and a population with known B-CLL sub-type with poor prognosis, thereby determining said discriminating value which identifies the B-CLL sub-type population having a poor prognosis.
- 30 4. The method according to claim 3, wherein the individual is a member of an unselected population.
5. The method according claim 3, wherein the individual is a member of a population
35 already identified as having a B-CLL sub-type with a poor prognosis.
6. The method according to any of claims 3-5, wherein the determination is performed at several time points at intervals as part of a monitoring of a cancer patient after or during the treatment for primary cancer.
- 40 7. The method according to any of the preceding claims, wherein the expression product is a transcriptional product.

8. The method according to claim 7, wherein the at least one transcriptional product is selected from the group consisting of SEQ ID No 2, SEQ ID No 4, SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11.
- 5 9. The method according to claim 7, wherein said at least one transcriptional product comprise a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:15.
- 10 10. The method according to claim 7, wherein said at least one transcriptional product comprise a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:16.
11. The method according to claim 7, wherein said at least one transcriptional product comprise a nucleotide sequence spanning the junction between Exon-2 and Exon-3.
- 15 12. The method according to claim 11, wherein the nucleotide sequence spanning the junction between Exon-2 and Exon-3 is the last 20 nucleotides of the 3'-end of SEQ ID No:15 and the first 20 nucleotides of the 5'-end of SEQ ID No:16.
- 20 13. The method according to any one of the preceding claims, wherein the presence of at least one of the transcriptional product(s) indicates that the individual has a subtype of B-CLL associated with a poor prognosis.
- 25 14. The method according to any one of the preceding claims, wherein the presence or absence of the transcriptional product(s) is/are determined by a method selected from the group consisting of a nucleic acid hybridisation based technique and a PCR based technique.
- 30 15. The method according to any one of the preceding claims, wherein the biological sample is selected from the group comprising blood, serum, plasma, urine, saliva, lymph node biopsy, bone marrow, spinal liquid, spleen biopsy, and liver biopsy.
- 35 16. Use of a compound capable of eliminating transcription of at least one expression product comprising a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 for the treatment of B-CLL.
- 40 17. Use of a compound capable of eliminating transcription of at least one transcriptional product comprising a nucleotide sequence selected from the group consisting of SEQ ID SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 for manufacture of a medicament for the treatment of B-CLL.
18. A method of treating B-CLL comprising administering to an individual in need thereof a compound capable of eliminating at least one type of transcriptional product, said

transcriptional product comprising at least one nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

5 19. The method according to claim 18, wherein all types of the transcriptional product(s) are eliminated.

20. The method according to any one of claims 18 or 19, wherein the at least one transcriptional product is selected from the group consisting of SEQ ID No 2, SEQ ID No 4,
10 SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11.

21. The method according to any one of claims 18-20, wherein the elimination of the at least one transcriptional product is achieved by inhibiting the formation of the transcriptional product.

15

22. The method according to any one of claims 18-20, wherein the elimination of the at least one transcriptional product is achieved by destroying the transcriptional product.

23. The method according to claim 21, wherein the compound is an nucleotide capable of
20 inhibiting the transcription of a nucleic acid sequence encoding any transcriptional product comprising a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

25 24. The method according to claim 22, wherein the compound is an nucleotide capable of destroying any transcriptional product comprising a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18.

30 25. The method according 24, wherein the nucleotide is an si-RNA.

26. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

35 i) an amino acid sequence of SEQ ID NO: 3,

ii) an amino acid sequence having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3

40 ii) a fragment of SEQ ID NO:3 having at least 60% sequence identity compared to the full length sequence of SEQ ID NO:3.

27. An isolated polypeptide according to claim 26, said polypeptide having interleukin or cytokine activity

28. The isolated polypeptide according to any one of claims 26 or 27, which folds as a 4-helical cytokine.
- 5 29. Use of an isolated polypeptide according to any of claims 26-28 in a diagnostic method for a subtype of B-CLL having a poor prognosis.
30. The use according to claim 29, wherein the diagnostic method is based on immunological assays, such as FACS analysis, western blotting, RIA, ELISA,
10 Immunohistochemistry etc.
31. Use of an isolated polypeptide as defined in any of the claims 26-28 for the treatment of cancer.
- 15 32. Use of an isolated polypeptide as defined in any of the claims 26-28 for the preparation of a medicament for the treatment of cancer.
33. Use according to claims 31 or 32, wherein the cancer is B-CLL.
- 20 34. A method of immunisation of a patient in need thereof against B-CLL, wherein said immunisation generates an immune response in said patient which recognises a translational product of SEQ ID No 2, SEQ ID No 4, SEQ ID No 6, SEQ ID No 7, SEQ ID No 8, SEQ ID No 9, SEQ ID No 10 and SEQ ID No 11.
- 25 35. A method for producing an antibody with specificity against an isolated polypeptide as defined in any of the claims 26 to 28, said method comprising the steps of
- i) providing a host organism,
 - 30 ii) immunising said host organism with an isolated polypeptide as defined in any of the claims 26 to 28, or transfecting said host organism with an expression vector capable of directing the expression of an isolated polypeptide as defined in any of the claims 26 to 28,
 - 35 iii) obtaining said antibody.
36. An antibody obtainable by the method of claim 35.
37. An isolated polynucleotide selected from the group consisting of:
40
- i) a polynucleotide comprising SEQ ID NO:5
 - ii) a polynucleotide encoding a polypeptide having the amino acid sequence of SEQ ID No 3,

iii) a polynucleotide, the complementary strand of which hybridises, under stringent conditions, with a polynucleotide as defined in any of i) and ii), and encodes a polypeptide, which has at least 60% sequence identity with the amino acid sequence of SEQ ID No 3,

5

iv) a polynucleotide which is degenerate to the polynucleotide of iii), and

v) the complementary strand of any such polynucleotide.

10

38. The isolated polynucleotide according to claim 37, comprising the nucleotide sequence selected from the group consisting of SEQ ID No:2, SEQ ID No:4, SEQ ID No:6, SEQ ID No:7, SEQ ID No:8, SEQ ID No:9, SEQ ID No:10 and SEQ ID No:11.

15 39. A diagnostic kit for *ex vivo* or *in situ* diagnosis of a subtype of B-cell chronic lymphocytic leukaemia (B-CLL) in a individual, the kit comprising a detector molecule capable of detecting the presence or absence of at least one expression product, wherein said at least one expression product comprise a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15,
20 SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 in a biological sample isolated from the individual.

40. A kit according to claim 39, wherein the detector molecule is a nucleotide.

25 41. A kit according to claim 40, wherein the nucleotide is capable of hybridising to a nucleotide sequence selected from the group consisting of SEQ ID No:12, SEQ ID No:13, SEQ ID No:14, SEQ ID No:15, SEQ ID No:16, SEQ ID No:17 and SEQ ID No:18 under stringent condition.

1/13

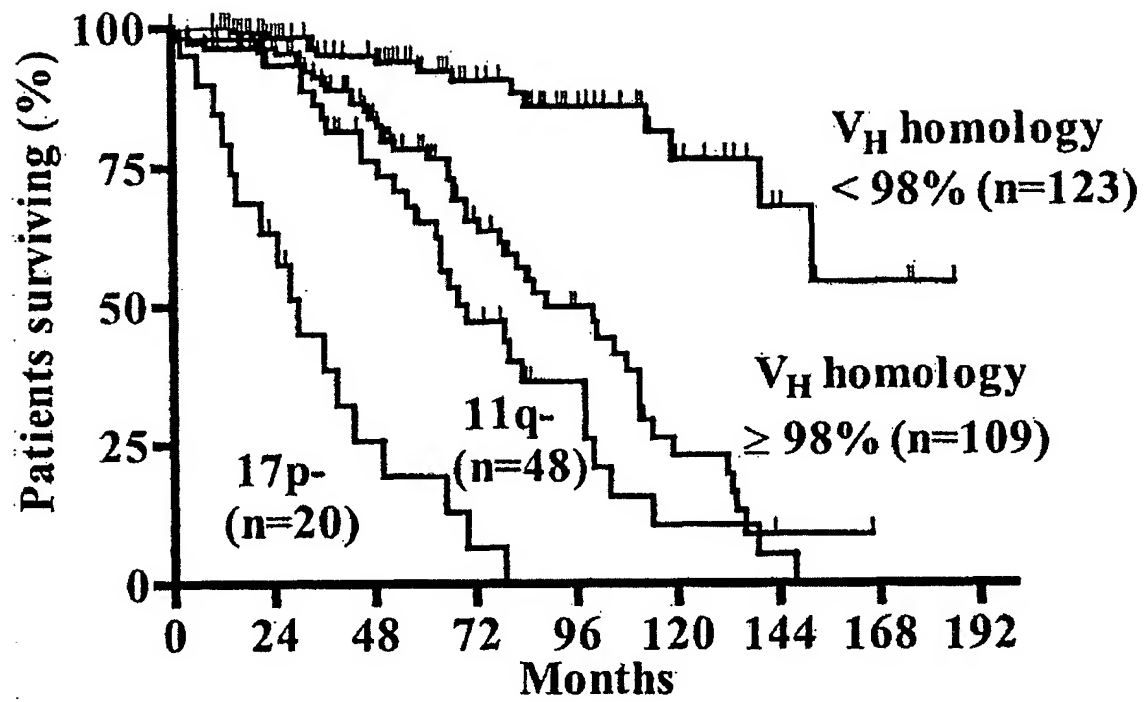


Fig. 1

2/13

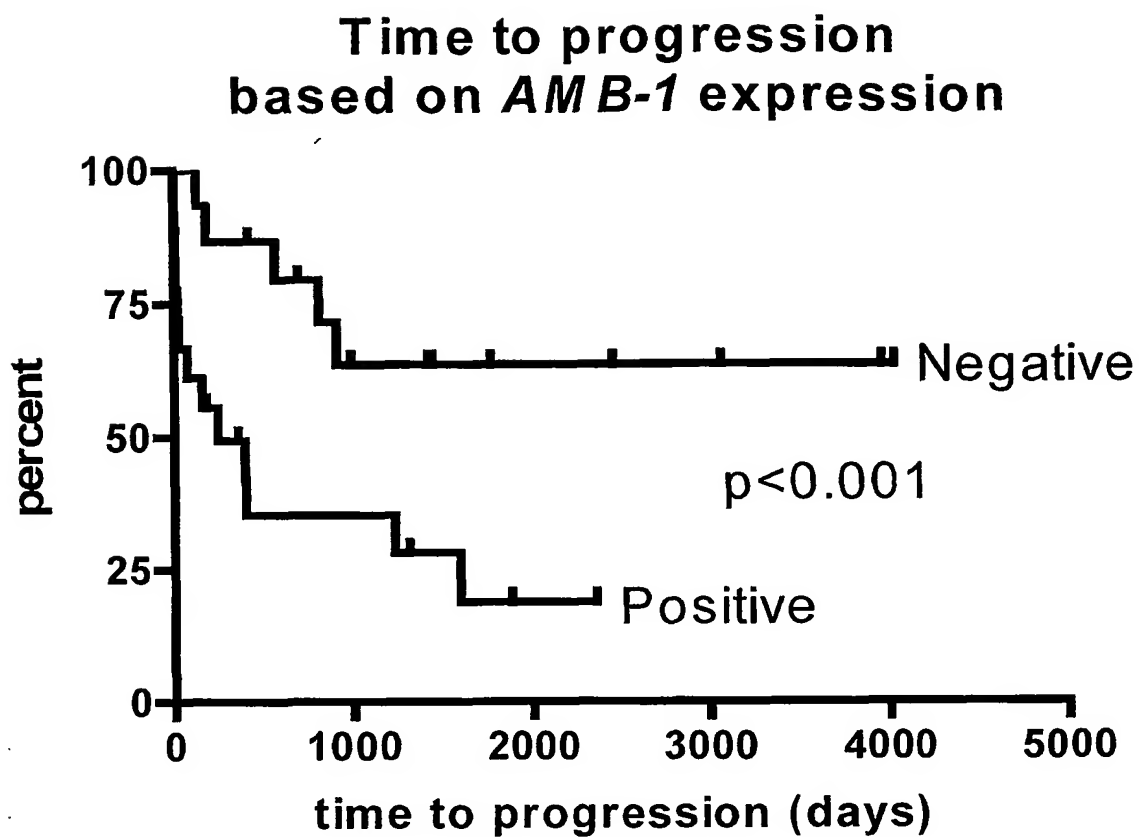


Fig. 2a

3/13

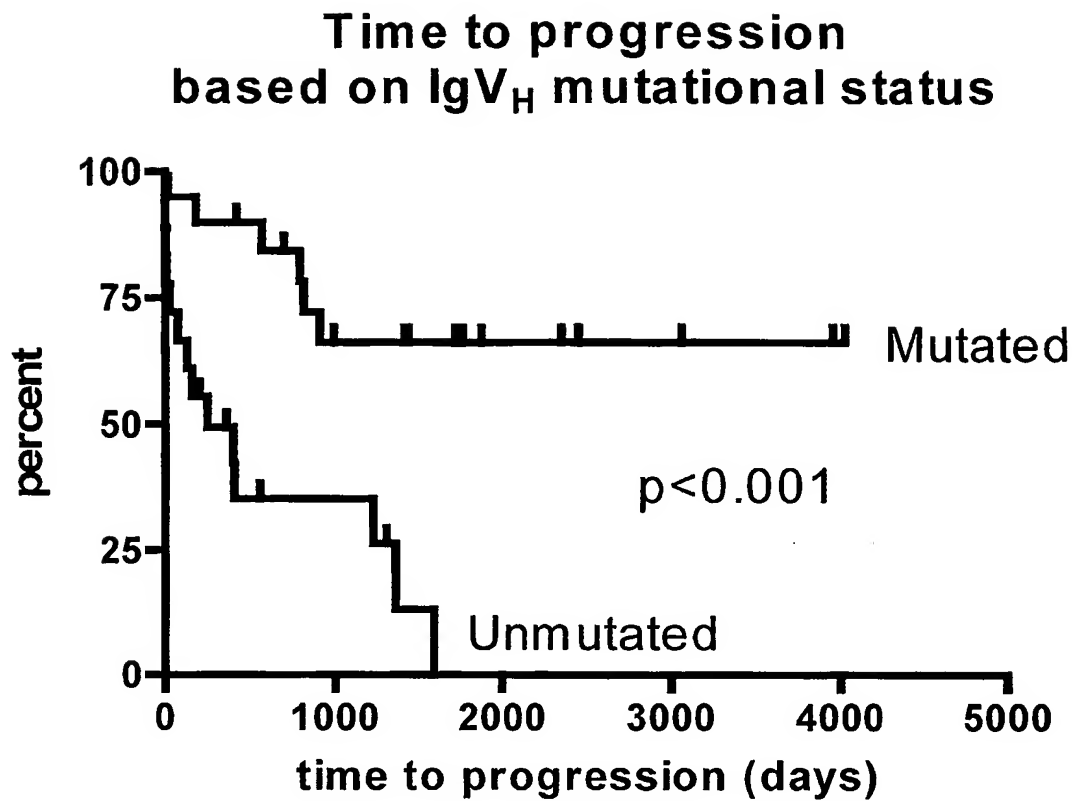


Fig. 2b

4/13

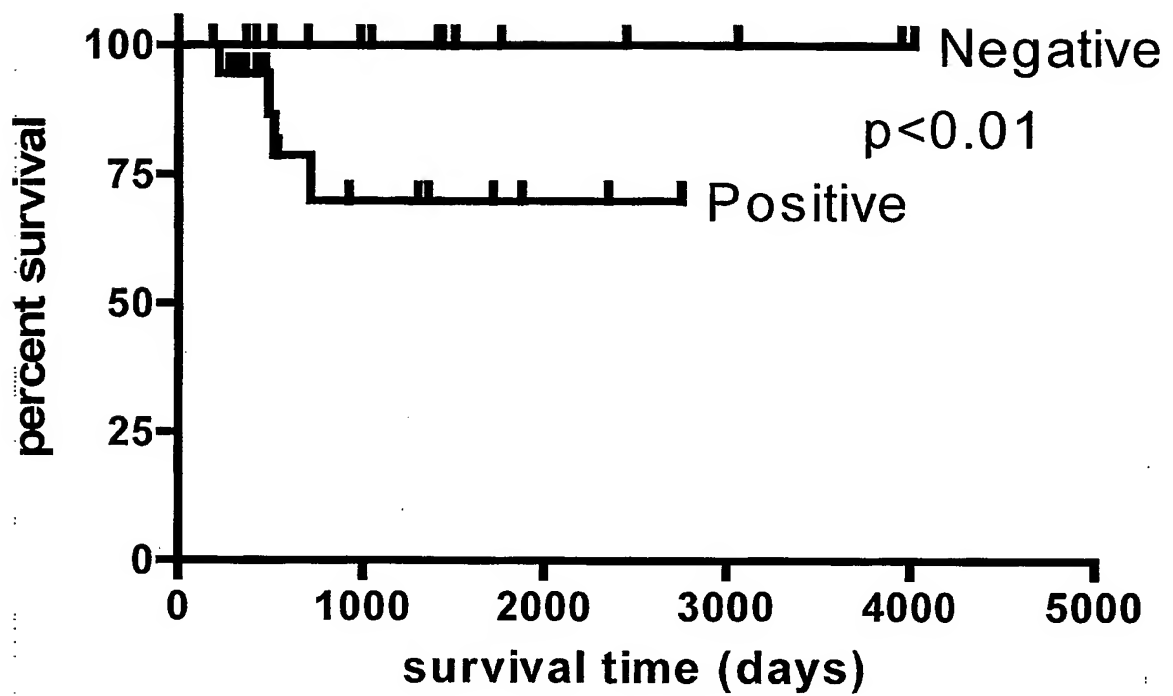
Survival based on *AMB-1* expression

Fig. 2c

5/13

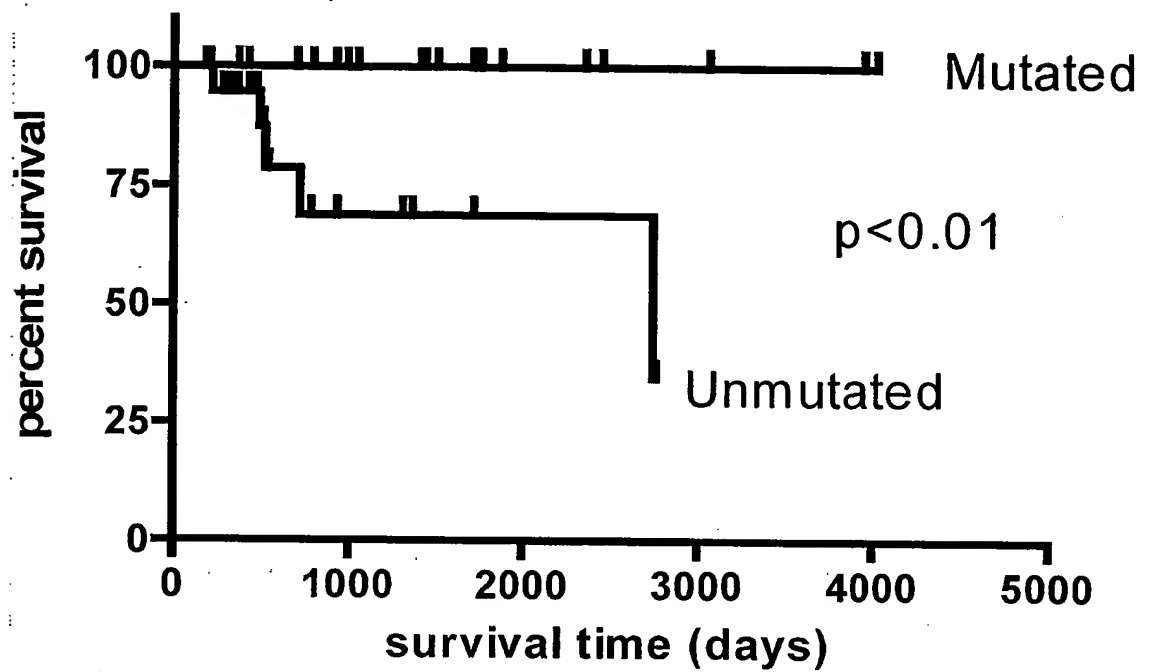
Survival based on IgV_H mutational status

Fig. 2d

6/13

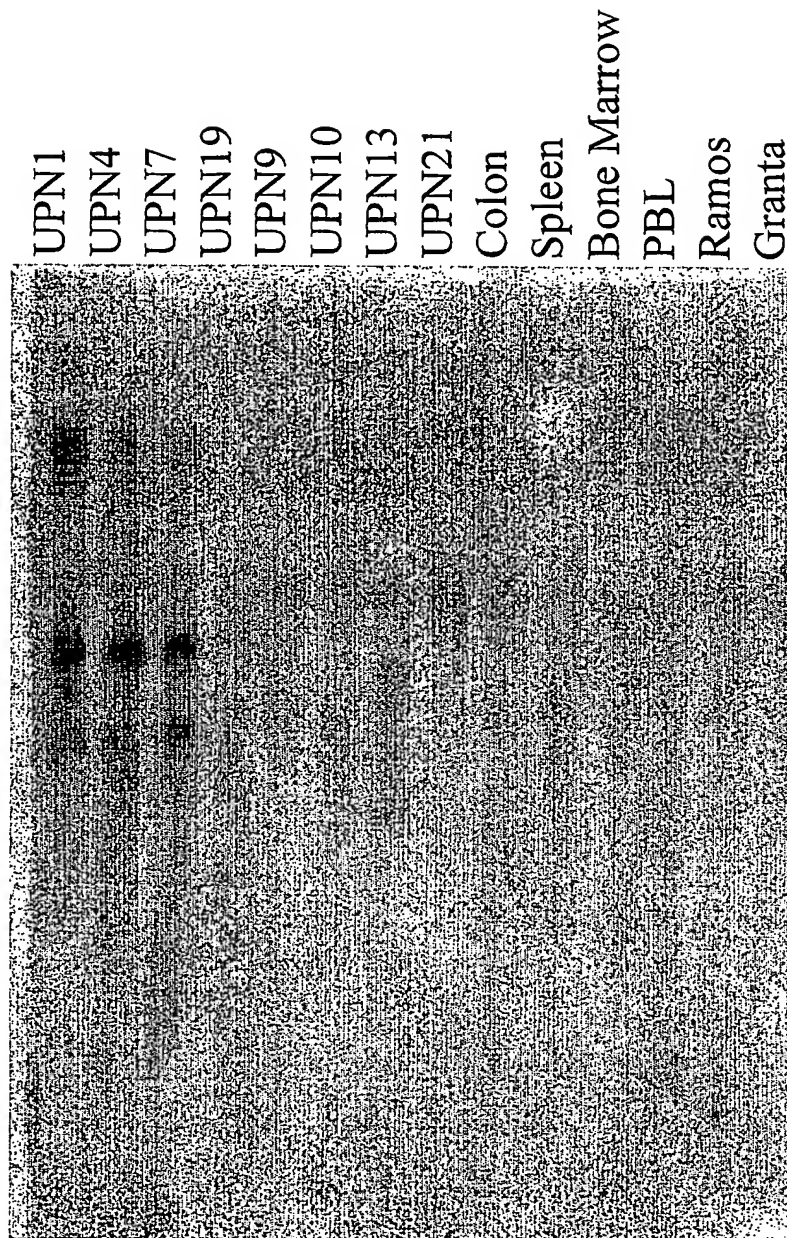


Fig. 3

AMB1_pe_pss	CCCCCCCCC	CCCCCCCCC	HHHHHHHHH	HHHEEECCC	CCCCCEEE
AMB1_pe_seq	MFNKCSEHSS	IYRPAADNSA	SSLCAIICFL	NLVI ECDLET	NSIINKLIY
-----		- - -	-+L--II--L	N--+	+-- +L--+
dliara_seq		HKCD	ITLQEIIKTL	NSL	KTLCTELTVT
dliara_ss		CCCC	HHHHHHHHH	HHH	CCCCCEEE
CORE		0000	0010040015	005	0000100041

AMB1_pe_pss	EECCCCCCH	HHHHHHHHH	HHHHCCCH	HHH	HHHH
AMB1_pe_seq	LFSQNNRIR	FSKLLKILF	YI SIFSYPE	LMC	EQYV
-----	+F++-	- + - - -	+S+-E	--C	+Q--+
dliara_seq	DIFAASKNTT	EKETFCRAAT	VLRFYSHHE	KDTRCLGATA	QQFHRHKQLI
dliara_ss	CCCCCCCCC	HHHHHHHHH	HHHHHHHHH	CCHHCCCCCH	HHHHHHHHH
CORE	0430000000	0001200920	4900300000	0000000000	0000000030

AMB1_pe_pss	HHHC	CCCCCCC	CCCC	HHH	CCCCCE
AMB1_pe_seq	TFIK	PGIHYGQ	VSKKH I IYS	TFL	SKNFKF
-----	-F+k	-G+ + - - -	V + - + - - -	FL	SK--
dliara_seq	RFLKRLDRNL	WGLAGLNSCP	VKEANQSTLE	NFLERLKTIM	REKYSKCSS
dliara_ss	HHHHHHHHH	HHHCCCCC	CCCCCEEH	HHHHHHHHH	HHHHCCCCC
CORE	0160080008	0011000000	2000002060	0560080013	0000000000

AMB1_pe_pss	EEEEEC				
AMB1_pe_seq	QLLRVCW				

dliara_seq					
dliara_ss					
CORE					

Fig.4

8/13





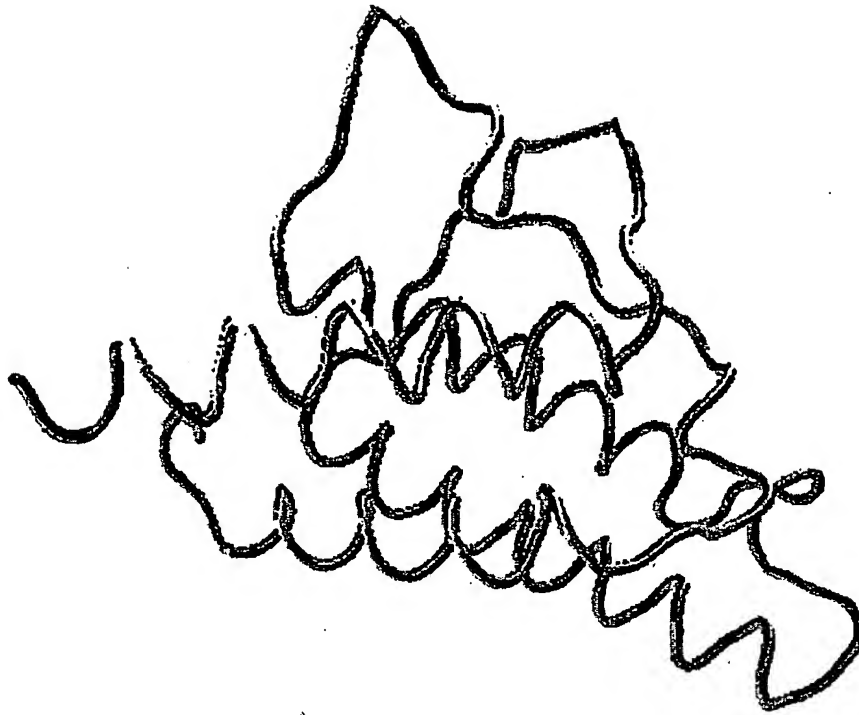
Fold Library	Template Length	Model	PSI-M E-value	SW-MED E-value	Alotext	Class	Fold	Superfamily	Family	Protein
d1h3212 23%id.	75		2.88	1	n/a	Alpha and beta proteins (ab)	Thioredoxin fold	Thioredoxin-like	Glutathione S-transferases, N-terminal domain	Glutathione S-transferase
d1lra_17%id.	129		3.05	1	n/a	All alpha proteins	4-helical cytokines	4-helical cytokines	Short-chain cytokines	Interleukin-4 (IL-4)
e1dga_17%id.	89		5.34	1	n/a	not in SCOP 1.53	PDB header: hydrolase inhibitor.	Chain A: PDB Molecule: chetg (protease inhibitor).		PDBTitle: solution structure of interleukin-2 1beta generation.
d1am2a_22%id.	80		5.57	1	n/a	All alpha proteins	Helix-loop-helix DNA-binding domain	Helix-loop-helix DNA-binding domain	Helix-loop-helix DNA-binding domain	SREBP-1a

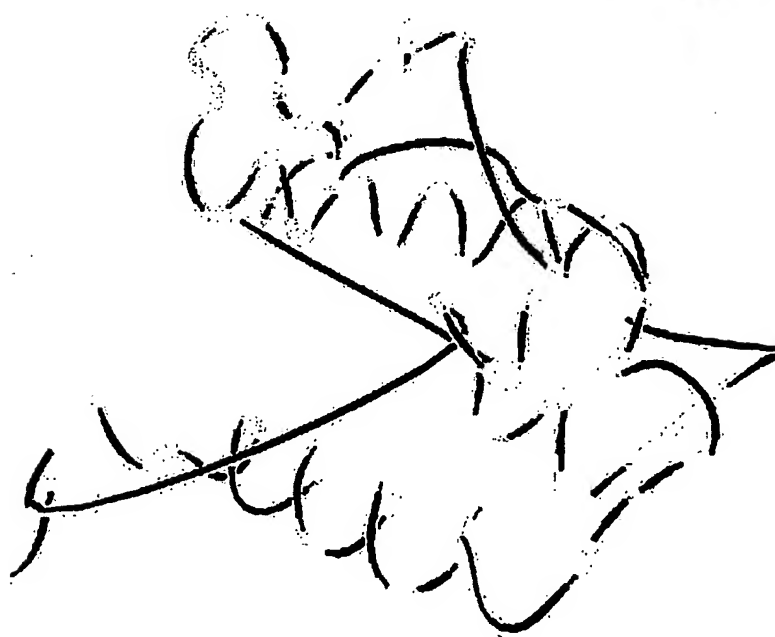
Fig.5

BEST AVAILABLE COPY

9/13



Human IL4



AMB1

Fig.6

10/13

AMB1 IL4 IL3 IL13 GM-CSF
 -----MENKCSFHSIYRPADNSASSICALICF-----LNVIECDLE-TNS-EINKLIY-LFSONRIRFSKLLKILFYI-SIFSYPE--IMC-----EQVTFIK-----PGHYGVSKKH-IIYS--TFLSKNFKOLLRVGV
 IL4 -----HKCDITIQEIKT-----LNSL-----TE-OKT-LCTELVTDFEASKNTTEKTCFRAATVLRQFYSHHEKTRC--IGATAQOFRRKOLIRFKRLDRLNGL-----AGINSCPKEANOSTLE--NFELRLATIMREKYSKCS-
 IL3 -----ANCSIMIDE-----IHH-----LKRPNPLDEN-----NLN-SEMDIEMERILATFNLIARVAVKHLENASIES-----ILKNLLSCLPL--ATAAPTRHPHIKOG--DWNEL-----RRKLTFTYKLTENAOAQ
 IL13 MAULLITVIAUTCIGGFASFGVPPSTALRELIEE--LVNI-----TONQKAPLCSNMWSI-----NLTA-GMYCALESINV-----TORMLSGTCPH-KVSAGOFSSLHVRDTKIEVAQE-----VOLLHLKGLREGREN
 CSF -----RSPSPSTQMEHVNVAQEARLLNL-----SRDPAENETVEVISEM--FDLQE-PTCLQTRLELKHQ-----GLRSLTK-----LGPITPMASHYKQHCPTPTSCATQIITTESF--KENLKOFILNITPDONEP

Fig. 7

11/13

	1	2	3	4	5	6	7	8	9	10	11	12
A	whole brain	cerebellum left	substantia nigra	heart	esophagus	colon, transverse	kidney	lung	liver	leukemia, HL-60	fetal brain	yeast total RNA
B	cerebral cortex	cerebellum right	accumbens nucleus	aorta	stomach	colon, descending	skeletal muscle	placenta	pancreas	HeLa S3	fetal heart	yeast tRNA
C	frontal lobe	corpus callosum	thalamus	atrium, left	duodenum	rectum	spleen	bladder	adrenal gland	leukemia K-562	fetal kidney	<i>E. coli</i> rRNA
D	parietal lobe	amygdala	pituitary gland	atrium, right	jejunum		thymus	uterus	thyroid gland	leukemia, MOLT-4	fetal liver	<i>E. coli</i> DNA
E	occipital lobe	caudate nucleus	spinal cord	ventricle left	ileum		peripheral blood leukocyte	prostate	salivary gland	Burkitt's lymphoma, Raji	fetal spleen	Poly r(A)
F	temporal lobe	hippocampus		ventricle right	ileocecum		lymph node	testis	mammary gland	Burkitt's lymphoma, Daudi	fetal thymus	human Cot-1 DNA
G	p.g.* of cerebral cortex	medulla oblongata		inter-ventricular septum	appendix		bone marrow	ovary		colorectal adenocarcinoma SW480	fetal lung	human DNA 100 ng
H	pons	putamen		apex of the heart	colon, ascending		trachea			lung carcinoma A549		human DNA 500 ng

Fig. 8

12/13

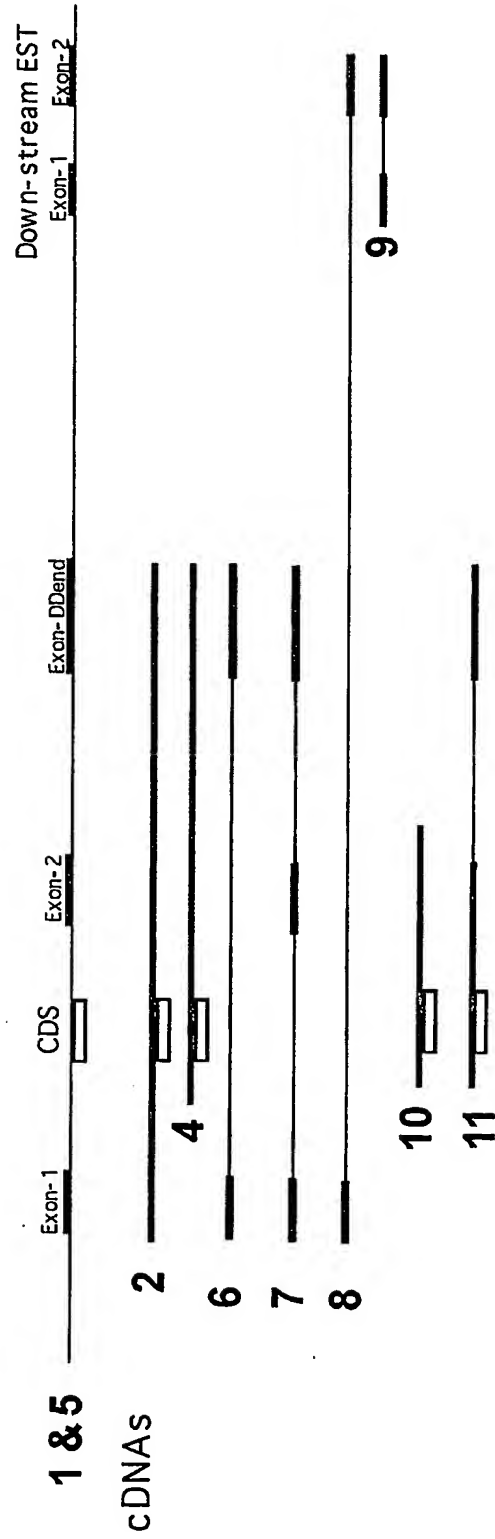
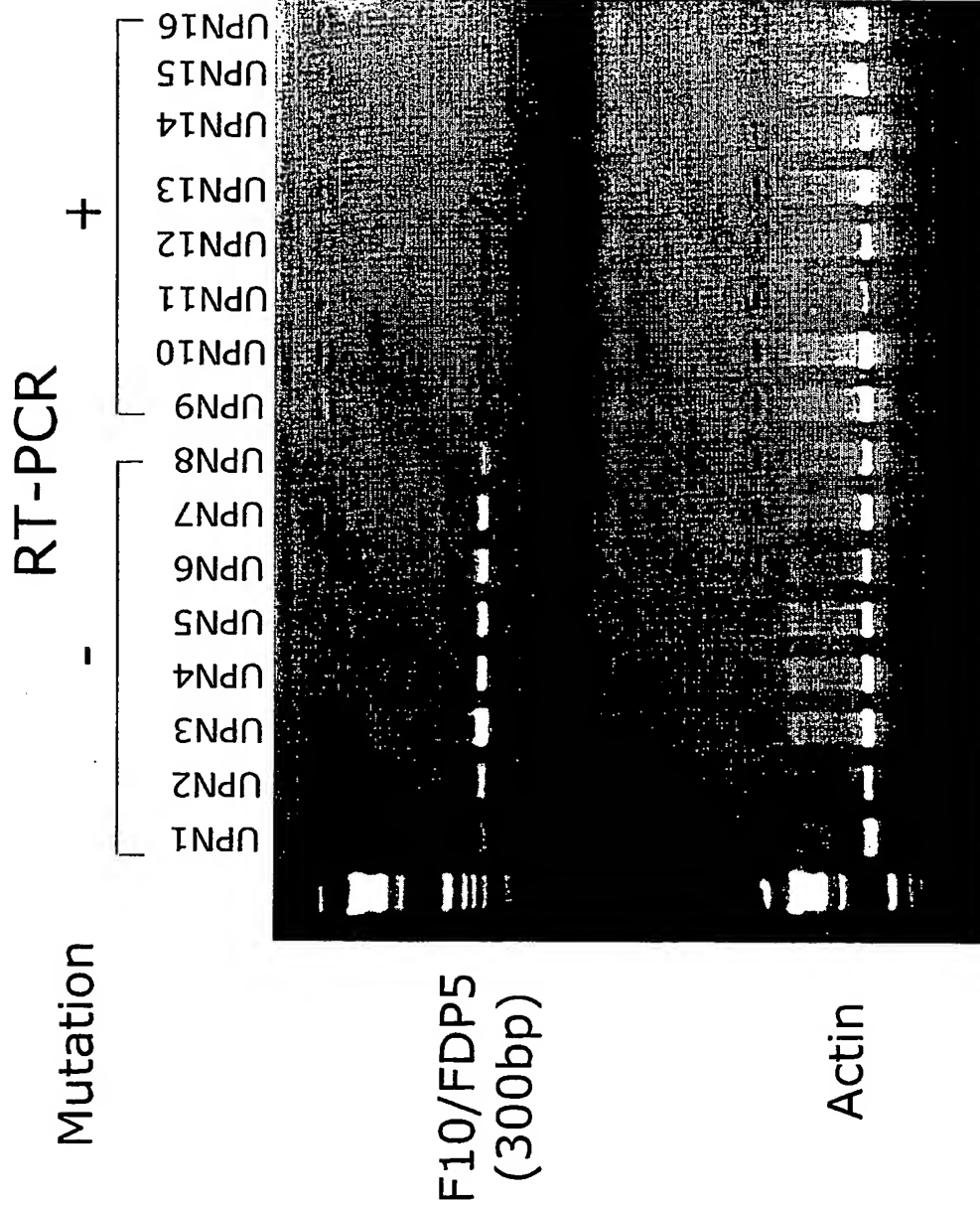


Fig. 9



OT Fig. 10 BEST AVAILABLE COPY

SEQUENCE LISTING

<110> Rigshospitalet
Henrik Leffers
Anne Mette Buhl Hertz
Jørgen Kjems

<120> Methods and kits for diagnosing and
treating B-cell Chronic lymphocytic leukemia (B-CLL)

<130> P34546PC01

<150> DK/PA 200201792

<151> 2002-11-19

<160> 43

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 19999

<212> DNA

<213> Homo Sapiens

<220>

<221> gene

<222> 40000-60000

<223> Sequence of ac063949.emhum

<400> 1

```
cacacgtagg ctacgagtgg ccctcagcct gcctcatcat ggacctgtgt tataataaat 60
atgtttaatt gtgctgtttt cttatagagg aaagtcctga tgttagttgc cttgaagtca 120
gacacccaga gagaatcaca ggttttcaga ttaattcatc gcttgattct tatccctgaa 180
gtcatatctc tggatctctg gttctcacat tataaatttc aatgattctt tttctatatg 240
gccatgtcat tcatatcctg tgtaatatgg ggaaactgag gtatgaatga catcattcaa 300
aaagcacctg caatTTTTtT ttgccaaagca cttacagctt tttctcatgt tgctttcaaa 360
aagtcattga aatattgttc acatattttg cagatgagga aatgaatatt caaatgcatt 420
aggtatcttg tccaagttct tacagccaga aagtagagaa atgaatttga attacaaatc 480
ttctacctct tggcttatgc tcttttcatg acactgggaa taaatgtctg aacaagcatg 540
acttcatgtt tcaactatTT atcaaatact tgttttctac taagatcttg cactcactca 600
gtgggatccc ctgaagcctg ctgattatTT gtcctttggc atttatcact ctctgtggga 660
ccttactctc ctatggtaaa gttttattgt tattaanaag attatttgac aataaatgta 720
gaaatcctac agatcatact caacaacatg tctaattgtc gcacacaatg tctaacaatc 780
atTTtatgaat actTTtatgtc aaacataagc aataaccta ttaaggaagg tatttttaat 840
aaattgacac tttttgacat aaccatattt caagtggctc cattgttttg tttatttatt 900
tattttattt tttatttatt tatttttgag aaagggtctc actctgttgc ccagactgga 960
gtgcagtggc aacatcatag ctactacag cctcgacctc tctgggctca agcaatcctc 1020
ccatctcagc ttcccaagta gctgggacta cagggtgtga ccatcatgcc aggctaattt 1080
ttcgatattt gtagagacgg ggttttgcct ggtcgtccgg gttggtctca aactcctggg 1140
tgttccgccc accttggcct cccaaagtgc tgggattata ggcatgagcc tcaagtggct 1200
actttttagg gttgaaattt atattgactg tcaactagct tccctagtta gtatttggga 1260
tctgctaact aatttatatt accatccaac ttgtcaacat ttgttgaaat ataactgtcc 1320
tcaacttttt tgtgtgaaca ttgaatacac tttcagacta aatttggttt attacttaat 1380
gtcttattct ttattagagt taataatatt tcttaatact ttgccttcca caaatgaata 1440
acttgtttgt gatggctacc tcttttttTt tcttagcctg tcacaggtat tatggataaa 1500
aattagcagc gctgggcaaa aacaatgaaa gaaatacact tgcctgggaa agctggggag 1560
gggtaaatga atataattca aaataccata tatttattca aactgttgg aatatatgtc 1620
ctgttggaat tgtaaaagtg acatatgttc tcttctggg tctcagactt ttaggatcta 1680
gttgagggaa ctggacttat acacaaaata caattcaaca acattatgag ctagaaaatc 1740
```

catgagctaa	agtctttggc	aaagacatta	ggtaacatga	ggagtcagga	aaaggagaaa	1800
ttactgtggg	ctggaatggg	ctgggaacat	gagatggagg	aagtggcttg	ttactggaga	1860
aaggatgagg	ttcaaagaga	tgggaaaaaa	agaaagagag	aagaaagaaa	agaaatgagg	1920
aaaaacaagt	tgccagaaag	aacaaggaag	aatagaggca	ggtaagcagt	ggattttgcc	1980
ctaggggaagg	taataatact	agagacggca	gtttctaaca	ggccatgatg	aataagatac	2040
actttagccc	tcattggtac	gtgcagaaat	tcaaatttgg	aaattcaagc	ttacatgaca	2100
gtaaatatat	gttgggaaaa	aaataaccgg	taaacattta	catcagctct	ttttcctaaa	2160
gagaaacctt	ttccatgcta	tgaatatatt	gtcacaattc	tgttttcaaa	atacttgctc	2220
tacttttcca	agccacaaga	ggaacatttt	tctctgccaa	cactctctga	ccttaaccag	2280
tttctccact	acgtctactc	ttaaagctctc	tttagagctg	tgtgtatctc	gtctttatgt	2340
aaacctocta	atgatataac	ttatggaaat	attcaggcaa	ctttttcatg	aactttacca	2400
ggaaagacat	ttctagcagg	agagcatgaa	tagaaatgga	ctcttcccca	gtctctgctg	2460
ggttctgact	gtggtcactc	taactataaa	aagtgtgtaa	aatcatgag	cagattatgt	2520
catttccttg	gggtccctaa	aaatttcaag	gtatctgtat	tagcacagga	agatttaaat	2580
tgattttctca	acacattcag	atatcttatg	aactttatta	agataaaatt	cctccagcat	2640
tcagaaactc	atatattaca	gaataaaaaa	taaagcagaa	aattagtgtg	cctggctaaa	2700
aatgagagca	gggttctatt	tcatttttga	aagtcactaa	gacagtaata	ataccattaa	2760
tgataaaaatg	ttacatttag	tttaattatta	agtgtgtttt	tgtatgccag	ccacataata	2820
tatactttta	tatgtatcac	ctacatttct	agatgtaaat	gtgagggaa	tatagtagta	2880
tctacctcgt	atgattgctg	gatgatttaa	atgagctggt	gtctcaaaaa	cttggtatag	2940
aaagcagaaa	cttttagtta	ttaaagattct	tactattcca	atatttgaat	aaaacagtga	3000
cctgctaaga	aaccccaata	atattctgat	acatcaaaac	cttctggcat	tagatgtttc	3060
taatctaaca	tcttcatatt	aattttttta	tgttttgatt	atctacattc	agtagtgaat	3120
gtgtttctaa	acgtcggatg	catttttaac	taaagtgtgt	ttgtaccaca	ttttgacaa	3180
ttttgtttta	actatgattc	agcttataac	aaaacaaaac	aatgcattct	ctctccactg	3240
ttaataaggt	taatgaaaag	ttgacttatg	aaaaaaatcc	taatttatgc	acatttctcat	3300
tgttttccct	gctaaggata	ttagtacttg	acgattctgt	aacaaagaat	tatcatggga	3360
tgaaactttg	atgcaaatat	cttatcaata	caatgtgctt	gattttacct	agatgagatt	3420
tttcttttct	tctttctttt	ttgagacagg	gttttgctgt	gttaccacag	ctagcctcaa	3480
accctcggcc	cctggcctca	agtgatectc	tcacctctgc	ctcccaaagt	gctgggtatt	3540
acagatgtga	gtcactgaat	ccagcctcac	ttagtggct	ttcttagtga	attattttat	3600
ctggttctaa	aactttttga	taatactctc	aaatatttat	ggattttata	acataattta	3660
tggattacgt	agttatgaat	ttcataaatg	attttgtgat	attgccacag	atcatcacca	3720
ttatacagga	tgtataacat	aaccatgggt	taataatatt	tcataaacta	tagaccaaac	3780
aaagactggg	caggaccagg	gcacgcacgc	atttttatatg	tgtgggtgcct	attggaatat	3840
gccaggcctc	ctgtgaaaaa	aatcagtaag	tgcttatctc	ataggaccaa	cggcccaaca	3900
ttctgaagt	cactaccaca	ctttgcactt	atctccatgt	ggaaatagat	agccactggt	3960
gaattctggg	gagaactgaca	cgtctgaaat	ctctcagctt	cacaaccctt	attacagccc	4020
tcagagaatc	ttctcacata	gcgccaaaaca	acaactttag	gaagtgatgt	tcctagaatg	4080
aatcaatttc	taaaattaaa	agtgaaaaca	atgacaagga	gaagggaggg	tcagagagga	4140
aaggctgatg	ttactaaaag	acaaaagaca	gtataacctc	ttatgaggat	ggtccagaca	4200
ctcagggaaa	tgaggaaga	aataaaaagat	aggagtttga	accacactgt	gatggctaac	4260
tttatgtgtg	gacccgaccg	atctatggga	caccagata	gcttgtaaag	cactatttct	4320
gggtgcgtca	gtgaggggtg	ttttggaaga	gatcaacact	tgagttagta	gactgagtaa	4380
agcagatggg	cctcaccaat	gtgggtgcac	attgtttaat	ctgttgagtg	cctggataga	4440
caaaaaaggc	aaaagaaggg	tgaattccct	ttctcgtctt	aagctgggac	agccatcctt	4500
tcctaccctc	agacatgaga	ggtttgatt	cttggtcctt	tggtcccaag	ggctgacact	4560
ggtggccacc	tctgggttca	ggtctttggc	cccagattgt	aagttacacc	atcagcttcc	4620
ttggttcttg	ggccttgaga	ctcaagctaa	aatacactac	cagcttccct	tgttctctag	4680
tgtagggaca	gcaaatcatg	aaaccttctg	cctccataat	catataagtc	aattcccgtg	4740
ataaatctgt	gcttatatat	ctatagcttt	ccttttggtc	tgtttctcca	aagaacctta	4800
atgtacacac	tatatgacct	aacctgtagt	aatgataacc	ttatgcaggt	ttgaataaga	4860
tgatgggtatt	ctcagtatct	gggaggtatg	ggctagagtg	atgaaccacc	gccatgagcc	4920
taggactgag	gagatttctg	aaatgtggaa	tatttggtgt	caaaaccaag	agataatata	4980
gccatgtgga	aaacatgtag	aactatcgta	tgattcagca	acccaaccac	tgggaattta	5040
cccaaaggaa	aggaaaccag	tatattaaag	agaatctgca	ctcccatggt	tattgcagca	5100
ttattctcaa	tagcctagat	atggactcaa	cctaggagat	tagatgaatg	gacaaagaaa	5160
atgtggcata	tgtacaccat	gaaatactta	ccagctataa	aaaagaatta	gccaagcaga	5220
tggtgtgtgc	ctgtcatccc	agctccttgg	gaggctgagg	tgggaagatc	tcgaggccag	5280
aagtttgaga	ccagcctggg	caaaataata	agactcggtc	tctaaaacaa	tttaaaata	5340
ggccttccct	aaaaaaagaa	taaaatcatg	tcattcacgg	caacatagat	gggactggag	5400

gatattactg ttaagtga aa ttagccagga acagcaagtt aaacccaca tattctgatt 5460
catatgcgga agctaaaaaa acgttgatct catagaagta aaaagtagaa cagaggatgc 5520
tggagactag aaaaggtagg gagaaggaag ggagagggaa aaatttggtta acagggtaca 5580
aaacaaaatt acagtttagtt agggagaatt aattccagca tcctgtagca ctataggatg 5640
actatagtta ataataatac ttttaattagt ctcaaatagc tagaaggagg atattgaatg 5700
ttcccaacac acacaaaaaa atgataatgt atgagatgat ggatatggta gttatcctga 5760
tctgatcact ctacattata tgtatcaaca catcactatg taccacacaa atatgtagaa 5820
tttttatttg tccattttaa aaagataaca aatttaaaaa taaaataaaa actaaattag 5880
tgttccatgt aaacctggat gaactgggtca ccctacgtct gcccatctag atggctggtc 5940
aaagtttccc aggtccaca tcaagttggt ccactgctca ctggaacttc cctagtccag 6000
ttggggcaaat agtaatttac agcaatagtg aatttatcac tgacatttct tcagttcccc 6060
tctttggcat ctgcttcttc ttttctgtaa tgctgtttgt tgaaatgcc aacattcttt 6120
ttcttcccta gagctattca ggggtgacct tcttttcgca ttttcccatg ccacttccat 6180
tatatcaaaa taaaacagtc ctgtgtggcc actgctcatg acctgtttc ctgccatgtg 6240
aagataggat cggctgctct ttcttctcct cctttttttt cagagacagg atctctccct 6300
gtcaccagga ctggagtgc atggcacagt cgtagcttgc tgcagcctcg aactcctgga 6360
cctcctcagc ctctgagta gctgggacta cagggtgcaca ccaccatgcc ttcctaattc 6420
gatatatata tatatatata ttatatatat aaaaatatata taaaatatata tatattttat 6480
atattatata tataatatat atattttata tataaaatat atatatata tatatatatt 6540
atatataaaa tatatatatt atatatatat atatatatta tagagatggg gtcttgctct 6600
gtcaccagga ctgaagatca gctgctcttt ctaatctgtg gttagataag atctgtctcc 6660
caggggataa aatactacct ggaataaagg tatctttaa ataatcccag agaagaaaac 6720
atttttatag tatgacagag gcagagaaaa cagagaatat ttgttaaggc aggactttca 6780
ccactccag tacaatcac tgtctgttac ctgcatact tacacgggct ggcactgctg 6840
ggggtacaaa gttagtgcca aacttcacaa tgggttagatt catgtttaaa aagccattgg 6900
atcaaacctt tgtgaaagtt tccagctttt ttctgttcca aatatgtgtc cattataaaa 6960
gaatctcaag agcataattg ccaagatagt ctatgtccat gagtatttca acatctctca 7020
tgaaatctgt tcccatcatt actcaagata ttgtatgaac agtattccac ataaactagg 7080
tgctcaataa tgattgattg gccaatggag ggtcattatt taatgcacta caatctttta 7140
tgcaaggggc ccacaggaat cagtatgat ccataggaat ccttttcttt tccattgaaa 7200
aagaaacaga tagtggttg tattaggttt ctgtgtgtg ttgtgagggt gaaagatatg 7260
aaaagaaatt tgatcagagc ataaatctga gcccatggga taggaaagaa tgagggaata 7320
aggaagaaaa cacagattat agacaggaaa atcaaaccta ttaaaactga taattttcga 7380
atactaaaaa tgtacattca tttgaacaaa aagattctat aaagcaagat ttctctgttc 7440
ttaccagcac taccatgccc aaactacctt aggaaatgaa tagcagagtc aaacttaaaa 7500
gcacctgaaa tttaaaacaa aaaccaattt acattttatt taagaaaagc aaacagatgg 7560
gcctgctaac aatgtcaaag tctcgtttac aaagaaaaaa acaaatctgg aacctgaagt 7620
caaacagagtt caaaataaaa agcaaaccaa taaacagaaa ccaacataaa cagaagttac 7680
taccatctcc ctgagcctgt gaaattctgg aacttctctt tctttctcgc cttcttcttc 7740
tctcacctgg aagacgagca gagtgaacac atcaggggtt gtcagttccc cagatggcac 7800
cacattcata aaccaccgac tccaggagaa tgtaggaagc ttagttaagg ccaaagttct 7860
ctttggatct tcctcatggg cttcaaggca aaagaaaaaa aagtttgctt gagaatatct 7920
tcatactcat tagtttgaac catgcaaaat tacagttttt ataggtaaaa tgagtgcata 7980
ttggcaattt caaatgatta accctaatac attatgcttt tgggtataga aatattcaga 8040
tcttaaacat atgctgttac atacaaaatc aggtatattc ctgcttctat aattaaagca 8100
aagagaattt cttttggtca ctactccttc tgacatgagg tatgaaccaa gttcaggacc 8160
cctaaaggtc tgggtctggg tcatttctcc acctctaact tgtgccgctt tcttggtcag 8220
tcattgtgtt ctgagctgtc tcataaaaca tctgctatga ctttactttc tctgatagg 8280
gtggctttcc atcggtggca cttcggtggc cttattggta tgctttatac actggttctc 8340
gtttccaaat tggcattatt attgttatga ttctgtctgc tctccacat ttccatctt 8400
tctcctgatc tctctcacct gtacatttct tacattttct cctgtgcttc cttcttcca 8460
tcatcattgc ccaagtgtgt cttctttctt ctcttcttca catttctttt gccgctctc 8520
acatatgcag agatggctct tggttttcct tctgaaatct catagtttgg aggtaaaact 8580
gttagcaagg ccactgagaa' gagaacaaaa gggaaacata agagaaacca agtcactatc 8640
tctctcattt cctggtttct agaagtaaga cccaaagaac tcactgtttc agtgctttca 8700
gctcaggcca aactagggtg atcaaaactga gcttctgagt gctgatcaaa acctataaaa 8760
ccaagtagac agaccatcta caaatcttca ctgttaaata ccataaagaa tgaaaaggtc 8820
actaattggt aagactatat gtgtgataat taaatttatg catcaacctg gctaggctaa 8880
aggatgacca ggtagctggt aaaacattat tctgggtgtg tccataagag tgttttcgga 8940
agagatcagc atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca 9000
ggcatcatcc aatctgtcga aagcttgaat aaaacaaaaa gaggaaggga aaatttgctt 9060

```

cttttcttct tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag 9120
acttaaacca ggagttacac ctttggcttc cctgggtctc agttcctttg acttggactg 9180
aattacactg ccagggttcc tggttctcca gcttgcagat ggcagatcat gggacttctt 9240
ggcctccata attgtgtgag tcaatttcca ttttatttac atatccagtt atgcattgct 9300
taacaatgga gacaggttct gagaaatgca ttgttaagtg atttcatcat tgtgcaaaca 9360
tcatagagtg taactacaca aacctggaca gcatagacta ctacacatct aggctacatg 9420
gtgtagcttg taacctcatg ataagtatgt ataacatcat gataagtatg tatgtatcta 9480
ccatatctaa atgtagaaaa ggtacagtaa aaatatggta taatcttatg ggatcaccat 9540
cataatgca atccttttgta gactgaaatg tcattgtgta gtgcatagact gtatacgcac 9600
acatacacaa acacacacaa atatactatt ggttcttttt ctctgaagag ccctaataca 9660
atatgttata catttatatt gactctattt caaaatttat ggttttggtg aaacatatgt 9720
ggagatgggg catagtgtg ttgaaactggg ctgatgaatg ggtgggaggc 9780
atcatttggg acaagcccag ggcatcagct tatagatatc aagagctcaa caagagcact 9840
ttatggcaaa acctcccaca agacctctca gaagttgaga aactgctaaa agtttcttta 9900
tgacagatga catttatgga taaaataggg attagcagga ttctttaaat acttctgaac 9960
actaaccttc atttctacca ggcagtgggg ccccaagtgc agggccatag gaagtacaag 10020
tctgggagat actaggctgc actgtctgta gagaatctga aaaaaataa gactcactga 10080
aatgcagttt ggtataatta ttgccatgca tcataattct aaatcatact agtggcaca 10140
tactcttccc tgaaaaaaca ttttcttggg ttgaaattcta aataattggt gtggtcaca 10200
ctgagctttt aaatatataa atactttcaa gtttgcata ttttattacc tgttcttaa 10260
caaacattga attcaacatg aaaatgatta tgggaaacat tcgggtatac agtccctgac 10320
tcttaaggac tcaggtaaat acttagggtt tttcatggcc ctagtctttg ggttaccaca 10380
tgtttcttct tcaaatcaca gattcaaaat caagaatgat aacacagtga ttgtgtagac 10440
aaaataagtg aacccaaatt gcttgcttct gtcattctat ggaaccactg agagttttta 10500
cttgtgctta aaattttgaa tagtaaaaca gagtgtcaac ttcatgctgg aatatttttg 10560
gctttttaga cacaatttta agtacctgaa gtatttttac aagactaagt aacatcactg 10620
aaattacagc tttcttcttt ttaaaactgg tatttgttat aaaactaaag agcgaatcaa 10680
gaaaagcata attattactg attattacag gattattact gaaaaagaaa tgtacggaat 10740
agaggaggaa ggagttaaca aatgatccac tctgggtgtt gaaaaaccca ataagcctgc 10800
ttccaggaag tgcctaagac agagctggct cagcttgctg ggtcacagca tgtaaggaaa 10860
ctgctgggct acatgccacc atcctcagtt gtccagatag ataatcccat agccccatgg 10920
ggaaataatc ttttaattatg atatagctga caccattcaa agcactatgc taagtctttt 10980
atgtgaatta acttttgtca aatttatatt tcataaataa cccaaatatg tataccacta 11040
ttatcctacc ttaaagagga gaaactgagc tcctaaagtt taaatatcta acccaagtta 11100
agactgctag tcaccctagg ctattaactc aggcagtcta actcaggtat aataacatta 11160
tgctactgtt tgcagctttg actatgcctg aattataacg tcatgctatc taactaaaaa 11220
gctaagggaa ataaaatgag ccatagggct caatttcata aaaggagaga aaatactggg 11280
gaaaagtgat aatgcagagt ttaaaatatt tttgtaaaag tgccagagat tgagtataac 11340
aagtgtgacc aaaaaaaaaa aaaaaaaaaa aaaagggaaga aggtaaaaaa aagaggagg 11400
tctgagaaat agaaatatca gaggaaggaa ataaaggagg gtgagagtaa attctctttt 11460
agcattcaga ttccacagat tccacaaatc acatttcttt ttttaccac taaggaaaaa 11520
taacacttga cctaacattt cattgcagtt agctaaagga tgctagaaaa actatggttc 11580
agtgttttgc tctaatttct tcaggaatag agaaaagtga caaaaagatc agagaagaga 11640
agaaaggaaa ctatcagaaa aatacagaat tggagttaga tataacatat ttgggttgaa 11700
ggtaaaattt tatattgtaa tcttaagtat cttgctactt cagtttggtc cctggaacag 11760
cagcatcaga atctgccgag ggcttggtta aaaggcagaa tctcaggtcc catcccagac 11820
tactgaatc agaataataa tactgacaag atgccccggg attcatatgc acagtagagc 11880
tgggcaagtt ccattgtagc ctgtgattgt tttctgcaac ttagtatttc tgagttttcc 11940
caaggaaaga aacccaggcc ttagcttctg gcagacttgt gtttctcctt tacttactag 12000
ctgcatgact catgagcaag gaaatcaaac tttatgtgcc tgagtttctt catctataaa 12060
atggagacta taataatcat ctctaggct tggtttgagg atgttcaaca aatgctcctt 12120
tcattcctct atttacagac ctgccgcaga caattctgct agcagccttt gtgctattat 12180
ctgttttcta aacttagtaa ttgagtgtga tctggagact aactctgaaa taaataagct 12240
gattatttat ttattttctc aaaacaacag aatacgattt agcaaattac ttcttaagat 12300
attattttac atttctatat tctcctaccc tgagttagtg tgtgagcaat atgtcacttt 12360
cataaagcca ggtatacatt atggacaggt aagtaaaaaa catattattt attctacgtt 12420
tttgcctaaa aattttaaat ttcaactggt gcgctgtgtg tggtaatgta aaacaaactc 12480
agtacagtat tatttcagac agtatttaag cccctgtact taaacatatt cctcgtacca 12540
atgaagttac atgaaaagca aatttgtgtg agatatcgta gatggaagta aattagtctt 12600
tatgttcccc acaaattgaa atgcatttca aaaactctgt gtgtgtatgt gtgtgtgtga 12660
cagagtgtgt gtgagagaga gacagagaga tacgctttgg ttgcctccat aagctggctg 12720

```

ctatgattaa taagaccaag ttttctaaag aaaatgagat cataacaaaa gccctcttta 12780
tgactatctt ttatcagggg caaaaaggaa agagacaaaa cagcatgaaa tgatgagacc 12840
aagtgatgaa aattcattca caatgattgc tttcaagagt aatttctctt gggtaattca 12900
gcagcctggt acatggctc tctggagtga tagctaagt aaatgaagcc tctaaaagtg 12960
gattatcctg actaagaatat actcagccaa taatgcaaca gaaatccatt caaagcattc 13020
gggaaaaatt caaaagaata aatattcttt tttttttttt aaagttaatg acctacgac 13080
catttcttcc ctgactaaca agcagcaagc acttaaaaat atccagccag gatgaaatag 13140
aaaccacact gacttgtaa tatttttggt tgggtccagg gactcagatt ctaagccaaa 13200
ttctttgaat gatcttggca aatgtctcga attatttttg ccaacttttc tttatcttgg 13260
aaaaaaagtt tcatgaatgg gtgtcaaaat tgattagttt taaaaacctt tcttgcatg 13320
acgtatggca ccctaaact gtattagaaa aaagttaagt actctgtagt gtgaaaaatt 13380
cttaaggac accctctttt acaaaactcac aaaaacagcc tttggaatac ccacatgaag 13440
tagctgttgt tattgctttc tatataccta catcttgtct attataaaaa gactggtttt 13500
tggcaggtgt ggtggctcac acctgtaatt ccagcacttt gggaggccaa ggcgggcgga 13560
tcacctgaga tcaggagttc aggaccagcc tgatcaatat ggtgaaaccc agtctttact 13620
gaaaatacaa aaatcacccg ggtgtggtga cgggcgcctg tagtcccagc tactcgggta 13680
gctgaggcag gagaatcact tgaactcagg agtcagaggt tgcatgagc tgagatcatg 13740
ccactgcact ccagcctggg tgacagagca agactccatc tcaaaaaaaa aaaaaaaa 13800
gactggtttt tcaacagcta ttcccacccc tctgcatgga aatattcacc cagtcaattg 13860
ttttcctagt ttgggtaatg gccctctggg caggactgga gtggggcaca caggagaagc 13920
tgcaaaactat gtttagaagc atgtctggga aatgtcatgc aagaaaagac atatttaaag 13980
gtaggctttg catgaatgga aaaggagagt aattctatgt agagcagagc ctcttacttg 14040
cagttagaga agcaaaagtg gggaaagcaag aggaattatg cttttcatca gccaaatttg 14100
caggtaggag gattggctca gtcactttgg ctgaggtcga tgaaccagg tgtaaagaaa 14160
gtggactaga ttaatttcat ccattacagg aagaggagcc gtgaaagata atccagaaat 14220
cattgggatt tgatggtaga aggtattttg ggactattcc atttgaaatg agaaggtag 14280
tgacattctt tgaattcctt tcaagcaaag gattaaattt acccatgagt tgactcagaa 14340
aaaacataaa aagtattggt gctctgctca gagttttatc taactcattc tcaactctta 14400
ttccatgatg aaatgacata aatgaggttt tttattgttg ttgttgttgt tttctggaca 14460
caaggcaagg tagctacctg ggcagagctg ttttattttt ctatgccgtg gagagaaatt 14520
ggttaattgg ccattggaag cagtcattaa gatgttccca tgcgagtga ctttcaggg 14580
ttcccagctt ctgcatcctt ccctgtccct caattccatt gttggtgatg acaatgtctc 14640
tcccatcagc ctcatgaagt tctctctcat ttattaaaat ttgctttcag gaaaaatttt 14700
gaaaatgtgt ccagtaatgc ctgattggcc ccttatccta aaggcttaa ctggaggaag 14760
gaagctaaac tgagaaatct tgcaaatcat tgagccaaaa acgtattaat agcaagatct 14820
atcatttatt gactagtatg tggcaggcag tgccctttta tttaggcagg gagagttgat 14880
ggggggggcg ggggtcacac atcttaaaga ggtgctatct cctctatat aaatcatgta 14940
agtcaagaga gtaaggaaat gtctttgttt ggttatattc aggggattag agtatagat 15000
agaagatccc aagaaacctt gggatcattt tagactaaga aatgccaaat ccgccggcg 15060
cggtggctca cgcctgtaat ccagcactt tgagaggcag aggtggcgcg atcacaagg 15120
caggagattg agaccgtcct ggctaacgtg gtgaaaccc gtcttacta aaaatacaaa 15180
aaattagccg ggcgtggtg cgggcgcctg tagtcccagc tactcgggag gcggaggcag 15240
gagaatggtg tgaactcagg aggcggagct tgcatcagc cgagattgcc ccaatgcact 15300
ccagcctggg cgacagaacg agactccgtc tcagaacaaa acaaaaggaa atgccaatac 15360
cagcagaaat agagccaaat catgaacata agctaaacaa atgttggcag ttagacctag 15420
tggttaagag agcagactct taactagaac actgcactcc atgtcctcac ttagacct 15480
cactgtgggg ttctaattaa cccctgttac ttaccagtgg cagtcttaag gcattcctta 15540
agttcgttgt gcccatttt gttcatctgt agaaggggta ggatgacagt agtgtttact 15600
ttataggctt actgtgagca ttaaatgagt tactactgta tttgtaaagt gcttaaaatg 15660
ctgtccaaa agagtttggt aaacacttaa acttgcatct aaactgacag 15720
ctctcaataa ctggaaatga tcaagcatag gccctggaat ataagcagg ctacatgaag 15780
gcaaaaatgt tggtttcttt tgttcagccc tgtgcctaga tcaatatcta gtgatcatgc 15840
tcaagaaata ttgttgaatg aatcaatgaa cctaccgagg tagttacata aaagagttct 15900
gcatgagtac aaatctgggc aaagtaccc ccaaggaaat ttccactttt agattctgtg 15960
atttctttaa ggaactgata aattggtgtg atacaatgta aaaaaatgtg cctatatgat 16020
ttgagaaaaa cttattttct ctcctctttt tttcttctt tcttctctc ctccttctt 16080
tcttctctcc ctccttctt tcttctctc ctccttctt tcttctctt tcttctttt 16140
tttcttctt tcttctttt tttcttctt tcttctttt tttcttctt cttctctct 16200
tcttctttt cttcttctt ttgtgcctt ctttctttt tcttctttt tctgtcttt 16260
cttcttctt tcttctttt ctgcctttt ttctctttt ttttctttt tcttctttt 16320
tttcttttaa gcagaccatg tctgttagat gaatgcctt ttctagttaa aaggttaaac 16380

aggaaagtga agcacaatta tcaaggggtct ccagtcacat ccacatgttc ttaatcatta 16440
 tcttctttta cagtttcata tctccaggcc ttctattggg tcaggttggc atttcgctgc 16500
 cctttatgtg tgtgacaagt gaaaataagg aaagaaaaaa actcaagtga agaaaatcag 16560
 aatctgcgca gcagttcctg ggcgtttcag ctgcttccca catcacctgc ctcatcaagc 16620
 cccagcatcc atctccttgc tcatcttaca ccctgtgtgc atgacaggcc caccattcat 16680
 ttatcagagc aaaggctctc ccactattct ggttcacccc cctacttagc cagatataca 16740
 agaatatctg cacggatgac ctgcctcacc tgggagctca gaggagctca gattccatta 16800
 ctatcgacc aaggacagat ctcccagcaa gaatgacaga aaagactaac tgcccccata 16860
 atctcccttc caaaacacag ttctcttaat tctcccaga aaccagaatg tgactgctca 16920
 cctctctaag gacctgaaaa caactggcca ttctagctat ttaaatcaac tttaaaaaat 16980
 ccaaccgcca aaatattaaa caattttgtt tggaatgata acataactaa cctgctgaca 17040
 gctgcttctg ctagggtgcaa aaatggaaaa aaaaatactt ctaatcaggt caaatcactc 17100
 tacctttggg attctaaatt tactcatatt ctcaaagaaa tatattcagt catagtgggg 17160
 aaaataggat tattccttta gctcgataag caaccagaag ttcttccttc aaatcttgac 17220
 atttaataca tcagaaattg atttttggaa aactgtttcc tatgaagcta tctctgctg 17280
 aaggattttt cttttacaat ccagactata gaaggaaatt cacaacctgg actttcacct 17340
 ccattggctc gagttttact gaccaattcc cacctctgcc ttacacctaa cggaagtta 17400
 tgctgtttt ctcttcacat accccaacag ttacaaatgg ttgttattat taagcatctt 17460
 ttattttgtg gcctctgatt acatggctcc cttaaattttg acctaatcac aaaagattgg 17520
 taaaatttct taacatatta ataataattt gtttatgtgt caatatctta gcatgtatca 17580
 attaagacag aggtcttaac gttctctttt tgaaagagaa tattaggatt cagagatatt 17640
 aagagattct occaggatca cagttaggta acagagctgg attttagtcc aggtctgtct 17700
 acagctctaa cgtatataca ccctttgtat aacatgtcac gaattcagca taaagggatc 17760
 ttcagtgatc taagttaggg gtcagcaacc ttttctaaaa aggaccaa atgaatattt 17820
 caggctttgt ggacctatg gtctctatca taactgttca aatcacctag tagtgtaaaa 17880
 ggagccataa gcaaaatata aactaacgaa tgtggctgtt ttatgggatt ttttttaac 17940
 tctttattta caaaagcagg tggcagatca gaactcactt atgggccata gttctctgac 18000
 ccctgacctg agaaaatctt atatttatgg acaacattta gactgtgact tgccaagtaa 18060
 gaacaagaag ctctgtcaac tgaaggtaaa ggctggaagt ctgaaagcaa agagctgtct 18120
 ggtgttaatg ataagtgaag tagttaaagt tagaagatcc cagttataag aagcacaaa 18180
 aataatgacc atagactcct gaacaagaat gtctggactt ctggcttagg cactctgtt 18240
 gtatggtcca ggccaagtta cctaattctc ccaggcctcc atttcttat cattaaatga 18300
 agataataaa agtattttcc tcagagagct gtaagaataa actgagctaa cccatgtcaa 18360
 gcacatagaa tagggccag cctatattaa tttatcaata aatgccagct acatattagt 18420
 tctctatatt tttattcatt atcataaaat gtttatctac agattggcat tgtaaggatg 18480
 gagttaaaat tgtatgtatg tgaagggaat ttattcctgt tactattgat ctgcatcaca 18540
 ttaccacaaa tttagtggtc taaagtaaca acattcattt tgcaacaaa tttgaaattt 18600
 gaggagggtc tgtctgggaa gacttgtctc tgccctatgt ggtatcagca gggggaggct 18660
 tgacggactg gcacatgcc ttccagaatg gccactcgc atgcctgcca agttggtgct 18720
 ggctcttggc tgggagctca gctggggctg agtgctaggg tccttgggag gttccttgtg 18780
 gcctgaactt cctcacaca aggcgctgc ggtgcgagag tgagcatttc aagatagagc 18840
 caagatgaca ctgtattact gtgtaagacc cagcctggga attaatgtag cctcacttcc 18900
 atcccactct attttataaa agtgaattat taaggtcacc ccatattcaa ggggatagga 18960
 attagacttc atctgtatta agaaaaatgt ttttaaaaat tgtagacatg ttttaaaatt 19020
 ctaaagtcca ctactggct gcagattatt tatatataca tgcaagatac actcctacat 19080
 tctcttctta gaaggctcag ttgcaggta agatgaagct cttcaagtga gatttcttat 19140
 gtatttatcc tctcaatctg aagacttgta aactaagaga caagttattt gcaacctaca 19200
 tacgcaatat tcaatggtaa agtatacata ggacagccac tacagacact cttgttttaa 19260
 atagaggaaa atgagagcac ataacagtca ttggctcata gcaactctga tatccagaca 19320
 gcaaacacaa gcagctcttt ttttaggtct cagtcctact gcctggattc cctactgctc 19380
 ttgggtcttc cctccagggt cttggttctt ggacctctt tcatttaata ctatttctgt 19440
 tctttaagt tcaagctggc aaaatatgat tgtacaattc tgtttaaaat tccaggactt 19500
 cctgtgatcc ttattgggga atactccatt agacaagaat ctctttgaca taagccattc 19560
 tctacctgag atccctgtaa ggctgtgatg ggaccacata acctaaaaat tattagaaga 19620
 ctcatgttt actgagagaa tatgcctagc atatgcttag atccttagag gaactctgtt 19680
 tcaaagggtc tatgagacat taccttatat ctttctaagg tacaacaaa aggtcttttg 19740
 cttttgagtt tgatcttga gctgacacct tttcttaatt tgagaatccc ctgctctatg 19800
 gagagactga caaagagaaa tagttttata tttgaatgta acatcttgga tctttaatag 19860
 attatcttaa aattttcctg aaaatgtaac agttcctttt tttaaaattc attctcccta 19920
 cacacttatt atatatgact aaaagaaact ccctggcatt ttcaacattc tggtagaat 19980

ttttcttagc ccaatctac

19999

<210> 2

<211> 3893

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> short version of mRNA

<400> 2

atcagaggaa	ggaataaaag	gaggggtgaga	gtaaattctc	tttttagcatt	cagattccac	60
agattccaca	aatcacattt	cttttttttac	caactaagga	aaaataacac	ttgacctaac	120
atttcattgc	agttagctaa	aggatgctag	aaaaactatg	ttgcagtggg	ttgctcta	180
ttcttcagga	atagagaaaa	gtgacaaaaa	gatcagagaa	gagaagaaaag	gaaactatca	240
gaaaaataca	gaattggagt	aggatataac	atatttgggt	tgaaggtaaa	attttatatt	300
gtaattctaa	gtatcttgct	acttcagttt	ggccctgga	acagcagcat	cagaatctgc	360
cgagggcttg	ttaaaaaggc	agaatctcag	gtcccatccc	agactcactg	aatcagaata	420
taaatactga	caagatgccc	cgggattcat	atgcacagta	gagctggcga	agttccattg	480
tagcctgtga	ttgttttctg	caacttagta	tttctgagtt	ttccaagga	agaaaacca	540
ggccttagct	tctggcagac	ttgtgtttct	cctttactta	ctagctgcat	gactcatgag	600
caaggaaatc	aaactttatg	tgccctgagtt	tcctcatcta	taaaatggag	actataataa	660
tcctctccta	ggcttgtttt	gaggatgttc	aacaaatgct	cctttcattc	ctctattttac	720
agacctgccg	cagacaattc	tgctagcagc	ctttgtgcta	ttatctgttt	tctaaactta	780
gtaattgagt	gtgatctgga	gactaactct	gaaataaata	agctgattat	ttattttatt	840
tctcaaaaca	acagaatacg	atttagcaaa	ttacttctta	agatattatt	ttacattttct	900
atattctcct	accctgagtt	gatgtgtgag	caatatgtca	ctttcataaa	gccaggata	960
cattatggac	aggttaagtaa	aaaacatatt	atttattcta	cgtttttgtc	caaaaatttt	1020
aaatttcaac	tggtgcgcgt	gtgttggtta	tgtaaaacaa	actcagtaca	gtagtattca	1080
gtacagtatt	taagcccctg	tacttaaaaca	tattcctcgt	accaatgaag	ttacatgaaa	1140
agcaaatgtg	tgtgagatat	cgtagatgga	agtaaatag	tctttatgtt	ccccacaaat	1200
tgaaatgcat	ttcaaaaact	ctgtgtgtgt	atgtgtgtgt	gtgacagagt	gtgtgtgaga	1260
gagagacaga	gagatacgct	ttggttgctt	ccataagctg	gctgctatga	ttaataagac	1320
caagttttct	aaagaaaatg	agatcataac	aaaagccctc	tttatgacta	tcttttatca	1380
ggggcaaaaa	ggaaagagac	aaaacagcat	gaaatgatga	gaccaagtga	tgaaaattca	1440
ttcacaatga	ttgctttcaa	gagtaatttc	tcttggttaa	ttcagcagcc	tgttactatg	1500
gctctctgga	gtgatagcta	atgtaaatga	agcctctaaa	agtggattat	cctgacaaga	1560
atatactcag	ccaataatgc	aacagaaatc	cattcaaagc	attcgggaaa	aattcaaaaag	1620
aataaatatt	cttttttttt	ttttaaagtt	aatgacctac	gatccatttc	ttccctgact	1680
aacaagcagc	aagcacttaa	aaatatccag	ccaggatgaa	atagaaaccc	acctgacttg	1740
ttaatatatt	tgtttggtcc	cagggactca	gattctaagc	caaattcttt	gaatgatctt	1800
ggcaaatgtc	tcgaattatt	tttgccaaact	tttctttatc	ttggaaaaaa	agtttcatga	1860
atgggtgtca	aaattgatta	gttttaaaaa	cctttcttgc	agatacgtat	ggcacccctaa	1920
aactgtatta	gaaaaaaatt	tcatatctcc	aggcctttca	ttgggtcagg	ttggcatttc	1980
gctgcccttt	atgtgtgtga	caagtgaaaa	taaggaaaga	aaaaaactca	agtgaagaaa	2040
atcagaatct	gcgcagcagt	tcctgggcgt	ttcagctgct	tcccacatca	cctgcctcat	2100
caagcccagc	catccatctc	cttgctcatc	ttacaccctg	tgtgcatgac	aggcccacca	2160
ttcatttata	agagcaaagg	ctctcccaact	attctgggtc	acccccctac	ttagccagat	2220
atacaagaat	atctgcacgg	atgacctgcc	tcacctggga	gctcagagga	gctcagattc	2280
cattactatc	gcaccaagga	cagatctccc	agcaagaatg	acagaaaaga	ctaactgccc	2340
ccaaaaatct	ccttccaaaa	cacagttctc	tttaattctcc	caagaaacca	gaatgtgact	2400
gctcacctct	ctaaggacct	gaaaacaact	ggccatttca	gctattttaa	tcaactttaa	2460
aaaatccaac	cgccaaaata	ttaaaccatt	ttggttgga	tgataacata	actaacctgc	2520
tgacagctgc	ttctgctagg	tgcaaaaaatg	gaaaaaaaaa	tacttcta	caggtcaaat	2580
cactctacct	ttgggattct	aaatttactc	atatttctca	agaaatatat	tcagtcatag	2640
tggggaaaat	aggattatct	ctttagctcg	ataagcaacc	agaagttctt	ccttcaaatc	2700
ttgacattta	atcaatcaga	aattgatttt	tgaaaaactg	tttcctatga	agctatctct	2760

```

gcctgaagga tttttctttt acaatccaga ctatagaagg aaattcacaa cctggacttt 2820
cacctccatt ggtcagagtt ttactgacca attcccacct ctgccttaca cctaacggaa 2880
gtttatgcct gttttctctt cacatacccc aacagttaca aatggttggt attattaagc 2940
atcttttatt ttgtggcctc tgattacatg gtcccctaaa ttttgaccta atcacaaaag 3000
attggtaaaa tttcttaaca tattaataat attttgttta tgtgtcaata tcttagcatg 3060
tatcaattaa gacagagggtc ttaacgttct ctttttgaaa gagaatatta ggattcagag 3120
atattaagag attctcccag gatcacagtt aggtaacaga gctggatttt agtccagggtc 3180
tgtctacagc tctaactgat atacaccctt tgtataacat gtcacgaatt cagcataaag 3240
ggatcttcag tgatctaagt caggggtcag caaccttttc taaaaaggac caaatagtaa 3300
tatttcaggc tttgtggacc ctatggtctc tatcataact gttcaaata ccatgtagtg 3360
taaaaggagc cataagcaaa atataaacta acgaatgtgg ctgttttatg ggattttttt 3420
ttaactcttt atttacaata gcaggtggca gtcagaact cacttatggg ccatagttct 3480
ctgacccttg acctgagaaa atcttatatt tatggacaac atttagactg tgacttgcca 3540
agtaagaaca agaagctctg tcaactgaag gtcaaggctg gagttctgaa agcaaagagc 3600
tgtctggtgt taatgataag tgaaatagtt aaagttagaa gatcccagtt ataagaagca 3660
caaagaataa tgaccataga ctctgaaca agaatgtctg gacttctggc ttaggcactc 3720
ttgttgtatg gtccaggcca agttaccta tctctccagg cctccatttt cttatcatta 3780
aatgaagata ataaaagtat tttctcaga gagctgtaag aataaactga gctaaccat 3840
gtcaagcaca tagaataggg ccagcctat attaatatt caataaatgc cag 3893

```

<210> 3
 <211> 121
 <212> PRT
 <213> peptide

<220>
 <221> PEPTIDE
 <222> (0)...(0)

<400> 3
 Met Phe Asn Lys Cys Ser Phe His Ser Ser Ile Tyr Arg Pro Ala Ala
 1 5 10 15
 Asp Asn Ser Ala Ser Ser Leu Cys Ala Ile Ile Cys Phe Leu Asn Leu
 20 25 30
 Val Ile Glu Cys Asp Leu Glu Thr Asn Ser Glu Ile Asn Lys Leu Ile
 35 40 45
 Ile Tyr Leu Phe Ser Gln Asn Asn Arg Ile Arg Phe Ser Lys Leu Leu
 50 55 60
 Leu Lys Ile Leu Phe Tyr Ile Ser Ile Phe Ser Tyr Pro Glu Leu Met
 65 70 75 80
 Cys Glu Gln Tyr Val Thr Phe Ile Lys Pro Gly Ile His Tyr Gly Gln
 85 90 95
 Val Ser Lys Lys His Ile Ile Tyr Ser Thr Phe Leu Ser Lys Asn Phe
 100 105 110
 Lys Phe Gln Leu Leu Arg Val Cys Trp
 115 120

<210> 4
 <211> 6209
 <212> RNA
 <213> Homo sapiens

<220>
 <221> mRNA
 <222> (0)...(0)
 <223> long version of mRNA

<400> 4
 ggatgtgagt gggccttcag acttaaacca ggagttacac ctttggtctc cctggttctc 60
 agttcttttg acttggaactg aattacactg ccaggtttcc tggttctcca gcttgcatgat 120

```

ggcagatcat gggacttctt ggcctccata attgtgtgag tcaatttcca ttttatttac 180
atatccagtt atgcattgct taacaatgga gacaggttct gagaaatgca ttgttaagt 240
atttcatcat tgtgcaaaca tcatagagtg taactacaca aacctggaca gcatagacta 300
ctacacatct aggctacatg gtgtagcttg taacctcatg ataagtatgt ataacatcat 360
gataagtatg tatgtatcta ccatacttaa atgtagaaaa ggtacagtaa aaatatggta 420
taatcttatg ggatcaccat catatatgca atcctttgta gactgaaatg tcattgtgta 480
gtgcatgact gtatacgcac acatacaca acacacaca atatactatt ggttcttttt 540
ctctgaagag ccctaataca atatgttata cttttatatt gactctattt caaaatttat 600
ggttttggtg aaacatatgt ggagatgggg catagggtgt tgaactggga tagtgcctg 660
ctgatgaatg ggtgggaggg atcatttggg acaagcccag ggcacagct tatagatata 720
aagagctcaa caagagcact ttatggcaaa acctcccaca agacctctca gaagttgaga 780
aactgctaaa agtttcttta tgacagatga cttttatgga taaaataggg attagcagga 840
ttctttaaat actttcgaac actaaccttc atttctacca ggcagtgggg ccccaagtgc 900
agggccatag gaagtacaag tctgggagat actaggctgc actgtctgta gagaatctga 960
aaaaataata gagtcactga aatgcagttt ggtataatta ttgccatgca tcataattct 1020
aatcatact agtggctaaa tactcttccc tgaaaaaaca ttttcttggg ttgaattcta 1080
aataattgtt gtggtcacca ctgagctttt aaatatataa atactttcaa gtttgcata 1140
ttttattacc tgttcttaa caaacattga attcaacatg aaaatgatta tgggaaacat 1200
tcgggtatag agtccctgac tcttaaggac tcaggtaaat acttagggta tttcatggcc 1260
ctagtctttg gggtagcaca tgtttcttct tcaaatcaca gattcaaaat caagaatgat 1320
aacacagtga ttgtgtagac aaaataagt aacccaaatt gcttgcttct gtcattctat 1380
ggaaccactg agagttttta cttgtgctta aaattttgaa tagtaaaaca gagtgtcaac 1440
ttcatgctgg aatatttttg gctttttaga cacaatttta agtacatgaa gtatttttac 1500
aagactaagt aacatcactg aaattacagc tttcttcttt ttaaaactgg tatttgttat 1560
aaaactaaag agcgaatcaa gaaaagcata attattactg attattacag gattattact 1620
gaaaaagaaa tgtacggaat agaggaggaa ggagttaaca aatgatccac tctgggtgtt 1680
gaaaacacca ataagcctgc ttccaggaag tgcctaagac agagctggct cagcttgctg 1740
ggtcacagca tgaaggaaa ctgctgggct acatgccacc atcctcagtt gtccagatag 1800
ataatcccat agccccatgg ggaaataatc ttttaattatg atatagctga caccattcaa 1860
agcactatgc taagtccctt atgtgaatta acttttgtca aatttatttt tcataaataa 1920
cccaaataatg tataccata ttatcttacc tttaaagagg gaaactgagc tcctaaagtt 1980
taaatactta acccaagtta agactgctag tcaccctagg ctattaactc aggcagtcta 2040
actcaggtat aataacatta tgctactgtt tgcagctttg actatgcctg aattataacg 2100
tcatgctatc taactaaaaa gctaaggga ataaatgag ccatagggct caatttcata 2160
aaaggagaga aaatactggg gaaaagtgat aatgcagagt ttaaaatatt tttgtaaaag 2220
tgccagagat tgagtataac aagtgtgacc aaaaaaaaaa aaaaaaaaaa aaaaggaaga 2280
aggtaaaaaa aagaggagg tctgagaaat agaaatatca gaggaaggaa ataaaggagg 2340
gtgagatgaa atttctttt agcattcaga ttccacagat tccacaaatc acatttcttt 2400
ttttaccaac taaggaaaaa taacacttga cctaacattt cattgcagtt agctaaagga 2460
tgctagaaaa actatgttgc agtggtttgc tctaatttct tcaggaatag agaaaagtga 2520
caaaaagatc agagaagaga agaaaggaaa ctatcagaaa aatacagaat tggagtagga 2580
tataacatat ttgggttgaa ggtaaaattt tatattgtaa tcttaagtat cttgtacttt 2640
cagtttggtc cctggaacag cagcatcaga atctgccgag ggcttggtta aaaggcagaa 2700
tctcaggtcc catcccagac tcaactgaat agaataataa tactgacaag atgccccggg 2760
attcatatgc acagttaggc tggcgaagtt ccattgtagc ctgtgattgt tttctgcaac 2820
ttagtatttc tgagttttcc caaggaagaa aaccagggcc ttagcttctg gcagacttgt 2880
gtttctcctt tacttactag ctgcatgact catgagcaag gaaatcaaac tttatgtgcc 2940
tgagtttctt catctataaa atggagacta taataatcat ctctagggtt tgttttgagg 3000
atgttcaaca aatgctcctt tcattcctct atttacagac ctgccgcaga caattctgct 3060
agcagccttt gtgctattat ctgttttcta aacttagtaa ttgagtgtga tctggagact 3120
aactctgaaa taaataagct gattatttat ttattttctc aaaacaacag aatacgattt 3180
agcaaatatc ttcttaagat attattttac atttctatat tctctacccc tgagtgtatg 3240
tgtgagcaat atgtcacttt cataaagcca ggtatacatt atggacaggt aagtaaaaaa 3300
catattattt attctacgtt tttgtccaaa aattttaaat ttcaactgtt gcgcgtgtgt 3360
tggtaatgta aaacaaactc agtacagtag tattcagtac agtatttaag cccctgtact 3420
taaacaatatt cctcgtacca atgaagttac atgaaaagca aatttgtgtg agatatcgta 3480
gatggaagta aattagtctt tatgttcccc atcaattgaa atgcatttca aaaactctgt 3540
gtgtgtatgt gtgtgtgtga cagagtgtgt gtgagagaga gacagagaga tacgctttgg 3600
ttgcctccat aagctggctg ctatgattaa taagaccaag ttttctaaag aaaatgagat 3660
cataacaaaa gccctcttta tgactatctt ttatcagggg caaaaaggaa agagacaaaa 3720
cagcatgaaa tgatgagacc aagtgtgaa aattcattca caatgattgc tttcaagagt 3780

```

```

aatttctctt gggttaattca gcagcctggt actatggctc tctggagtga tagctaattgt 3840
aaatgaagcc tctaaaagtg gattatcctg acaagaatat actcagccaa taatgcaaca 3900
gaaatccatt caaagcattc gggaaaaatt caaaaagaata aatattcttt tttttttttt 3960
aaagttaatg acctacgatc cttttcttcc ctgactaaca agcagcaagc acttaaaaaat 4020
atccagccag gatgaaatag aaaccacact gacttggtta ttttttgtt tgggtcccagg 4080
gactcagatt ctaagccaaa ttctttgaat gatcttggca aatgtctcga attatttttg 4140
ccaacttttc tttatcttgg aaaaaaagt tcatgaatgg gtgtcaaaat tgattagttt 4200
taaaaacctt tcttgcagat acgtatggca ccctaaaact gtattagaaa aaaatttcat 4260
atctccaggc ctttcattgg gtcagggttg ctttctgctg ccctttatgt gtgtgacaag 4320
tgaaaataag gaaagaaaaa aactcaagtg aagaaaaatca gaatctgcgc agcagttcct 4380
gggctgttca gctgcttccc acatcacctg cctcatcaag ccccgagcat catctccttg 4440
ctcatcttac accctgtgtg catgacaggc ccaccattca tttatcagag caaaggctct 4500
cccactattc tggttcaccc ccctacttag ccagatatac aagaatatct gcacggatga 4560
cctgcctcac ctgggagctc agaggagctc agattccatt actatcgcac caaggacaga 4620
tctcccagca agaatgacag aaaagactaa ctgcccccaa aatctccctt ccaaaacaca 4680
gttctcttaa ttctcccaag aaaccagaat gtgactgctc acctctctaa ggacctgaaa 4740
acaactggcc atttcagcta tttaaatcaa ctttaaaaaa tccaaccgcc aaaatattaa 4800
accatttttg ttggaatgat aacataacta acctgctgac agctgcttct gctaggtgca 4860
aaaaatggaaa aaaaaatact tctaatacagg tcaaatcact ctacctttgg gattctaaat 4920
ttactcatat tctcaaagaa atatatctag tcatagtggg gaaaatagga ttattccttt 4980
agctcgataa gcaaccagaa gttcttctct caaatcttga catttaatca atcagaaatt 5040
gatttttgga aaactgtttc ctatgaagct atctctgcct gaaggatttt tcttttacia 5100
tccagactat agaaggaaat tcacaacctg gactttcacc tccattgggtc agagttttac 5160
tgaccaattc ccacctctgc cttacaccta acggaagttt atgcctgttt tctcttcaca 5220
tacccaaca gttacaaatg gttgttatta ttaagcatct tttattttgt ggcctctgat 5280
tacatggtcc cctaaatttt gacctaatac caaaagattg gtaaaatttc ttaacatatt 5340
aataatattt tgtttatgtg tcaatatctt agcatgtatc aattaagaca gaggtcttaa 5400
cgttctcttt ttgaaagaga atattaggat tcagagatat taagagattc tcccaggatc 5460
acagttaggt aacagagctg gatttttagtc caggctctgtc tacagctcta acgtatatac 5520
accttttgta taacatgtca cgaattcagc ataaagggat cttcagtgat ctaagtcagg 5580
ggtcagcaac cttttctaaa aaggaccaa tagtaatatt tcaggctttg tggaccctat 5640
ggtctctatc ataactgttc aaatcaccat gtagtgtaaa aggagccata agcaaaatat 5700
aaaactaacga atgtggctgt tttatgggat ttttttttaa ctctttattt acaaaagcag 5760
gtggcagatc agaactcact tatgggccat agttctctga cccctgacct gagaaaaatc 5820
tatatttatg gacaacattt agactgtgac ttgccaagta agaacaagaa gctctgtcaa 5880
ctgaagggtc aggctggagt tctgaaagca aagagctgtc tgggtgtaat gataagtga 5940
atagttaaag ttagaagatc ccagttataa gaagcacaaa gaataatgac catagactcc 6000
tgaacaagaa tgccttgact tctggcttag gcactctgtg tgtatggtcc aggccaagt 6060
acctaatctc tccaggcctc cattttctta tcattaaatg aagataataa aagtattttc 6120
ctcagagagc tgtaagaata aactgagcta acccatgtca agcacataga atagggccca 6180
gcctatatta atttatcaat aaatgccag 6209

```

<210> 5

<211> 80000

<212> DNA

<213> Homo sapiens

<220>

<221> gene

<222> (0)...(0)

<223> human genome sequence

<400> 5

```

cacacgtagg ctacgagtgg ccctcagcct gcctcatcat ggacctgtgt tataataaat 60
atgttttaatt gtgctgtttt cttatagagg aaagtcctga tgttagttgc cttgaagtca 120
gacaccagaa gagaatcaca ggttttcaga ttaattcatc gcttgattct tatccctgaa 180
gtcatatctc tggatctctg gttctcacat tataaatttc aatgattctt tttctatatg 240
gccatgtcat tcatatcctg tgtaatatgg ggaaactgag gtatgaatga catcattcaa 300
aaagcacctg caatttttct ttgccaagca cttacagctt tttctcatgt tgctttcaaa 360

```

```

aagtcattga aatattgttc acatattttg cagatgagga aatgaatatt caaatgcatt 420
aggtatcttg tccaagttct tacagccaga aagtagagaa atgaatttga attacaaatc 480
ttctacctct tggcttatgc tcttttcatg acactgggaa taaatgtctg aacaagcatg 540
acttcatgtt tcaactatct atcaaatact tgttttctac taagatcttg cactcactca 600
gtgggatccc ctgaagcctg ctgattatct gtcctttggc atttatcact ctctgtggga 660
ccttactctc ctatggtaaa gttttattgt tattaaaagt attatttgac aataaatgta 720
gaaatcctac agatcatact caacaacatg tctaattgtc gcacacaatg tctaacaatc 780
atztatgaat actttatgtc aaacataagc aataaccta ttaaggaagg tatttttaat 840
aaattgacac tttttgacat aaccatattt caagtggctc cattgttttg tttatttatt 900
tattttattt tttatttatt tttttttgag aaagggtctc actctgttgc ccagactgga 960
gtgcagtggc aacatcatag ctactacag cctcgacctc tctgggctca agcaatcctc 1020
ccatctcagc ttcccaagta gctgggacta cagggtgtga ccatcatgcc aggctaattt 1080
ttcgtatttt gtgagacagg ggttttgcct ggtcgtccgg gttggtctca aactcctggg 1140
tgttccgccc accttggcct cccaaagtgc tgggattata ggcatgagcc tcaagtggct 1200
actttttagg gttgaaatct atattgactg tcaactagct tccctagtta gtatttggga 1260
tctgctaact aatttatatt accatccaac ttgtcaacat ttgttgaat ataactgtcc 1320
tcactttttt tgtgtgaaca ttgaatacac tttcagacta aatttggttt attacttaat 1380
gtcttattct ttatttagagt taataatatt tcttaatact ttgccttcca caaatgaata 1440
acttgtttgt ctggctacc tcttttttct tcttagcctg tcacaggat tatggataaa 1500
aattgacacg gctgggcaaa aacaatgaaa gaaatacact tgcctgggaa agctggggag 1560
gggtaaatga atataattca aaataccata tttttattca acactgttgg aatatatgtc 1620
ctgttggaat tgtaaaagt acatatgttc tcttctggg tctcagactt ttaggatcta 1680
gttgagggaat ctggacttat acacaaaata caattcaaca acattatgag ctagaaaatc 1740
catgagctaa agtctttggc aaagacatta ggtaacatga ggagttagga aaaggagaaa 1800
ttactgtggg ctggaatggg ctgggaacat gagatggagg aagtggcttg ttactggaga 1860
aaggatgagg ttcaaaagaga tgggaaaaaa agaaagagag aagaaaagaa agaaatgagg 1920
aaaaacaagt tgccagaaag aacaaggaag aatagaggca ggtaagcagt ggattttgcc 1980
ctagggaagg taatataact agagacggca gtttctaaca ggccatgatg aataagatac 2040
acttttagccc tcatttgttac gtgcagaaat tcaaatttgg aaattcaagc ttacatgaca 2100
gtaaatatat gttgggaaaa aaataaccgg taaacattta catcagctct ttttctctaa 2160
gagaaacctt ttccatgcta tgaaatattt gtcacaattc tgttttcaaa atacttgctc 2220
tacttttcca atgcacaaga ggaacattt tctctgccaa cactctctga ccttaaccag 2280
tttctccact acgtctactc ttaagctctc ttttagagctg tgtgtatctc gtctttatgt 2340
aaacctccta gatgatatac ttatggaaat attcaggcaa ctttttcatg aactttacca 2400
ggaaagacat ttctagcagg agagcatgaa tagaaatgga ctcttcccca gtctctgtctg 2460
ggttctgact gtggtcactc taactataaa aagtgtgtaa aaatcatgag cagattattt 2520
catttccctg gggctcctaa aaatttcaag gtatctgtat tagcacagga agatttaaat 2580
tgattttcta acatctttag atattttatg aactttatta agataaattt cctccagcat 2640
tcagaaactc acatattaca gaataaaaaa taaagcagaa aattagtgtg cctggctaaa 2700
aatgagagca gggttctatt tcatttttga aagtcactaa gacagtaata ataccattaa 2760
tgataaaatg ttaacattag ttaattatta gatgtgtttt tgtatgccag ccacataata 2820
tatactttta tatgtatcac ctacatttct agatgtaaat gtgagggaat tatagttagta 2880
tctacctgt atgattgctg gatgatttaa atgagctgtt gtctcaaaaa cttggtatag 2940
aaagcagaaa ctttttagtta ttaagattct tactattcca atatttgaat aaaacagtga 3000
cctgctaaga aacccaata atattctgat acatcaaaac cttctggcat tagatgtttc 3060
taatctaaca tcttcatatt aattttttta tgttttgatt atctacattc agtagtgaat 3120
gtgtttctaa acgctggatg catttttaac taaatgtgtt ttgtaccaca ttttgacaac 3180
ttttgtttta actatgattc agcttataac aaaacaaaac aatgcatctt ctctccactg 3240
ttaataagg ttaataagg taatgaaaag ttgacttatg aaaaaaatcc taatttatgc acattctcat 3300
tgttttcctt gctaaggata ttagtacttg acgattctgt aacaaagaat tatcatggga 3360
tgaaactttg atgcaaatat cttatcaata caatgtgctt gattttacct agatgagatt 3420
tttcttttct tctttctttt ttgagacagg gttttgtgtt gttaccagg ctgacctcaa 3480
accttggcc cctggcctca agtgatctc tcacctctgc ctcccaaagt gctgggtatt 3540
acagatgtga gtcactgaat ccagcctcac ttagttggct ttcttagtga attattttat 3600
ctggttctaa aactttttga taatactctc aaatatttat ggattttata acataattta 3660
tggattacgt agttatgaat ttcataaatg attttgtgat attgccacag atcatcacca 3720
ttatacagga tgtataacat aacctaggtt taatatattt tcataaacta tagaccaaac 3780
aaagacttgt caggaccagg gcacgcatgc attttatatg tgtgtgtcct attggaatat 3840
gccaggcctc ctgtgaaaaa aatcagtaag tgcttatctc ataggacca cggcccaaca 3900
ttcctgaagt cactaccaca ctttgcactt atctccatgt ggaaatagat agccactgtt 3960
gaattctggt gagaacgaca cgtctgaaat ctctcagctt cacaaccctt attacagccc 4020

```

tcagagaatc	ttctcacata	gcgccaaca	acaactttag	gaagtgatgt	tcctagaatg	4080
aatcaatttc	taaaattaaa	agtgaataca	atgacaagga	gaagggaggg	tcagagagga	4140
aaggctgatg	ttactaaaag	acaaaagaca	gtataacctc	ttatgaggat	ggtccagaca	4200
ctcagggaaa	tgcaggaaga	aataaaaagat	aggagtttga	accacactgt	gatggctaac	4260
tttatgtgtg	gacctgaccg	atctatggga	caccagata	gcttgtaaag	cactatttct	4320
gggtgctgca	gtgaggggtg	ttttggaaga	gatcaacact	tgagtacagta	gactgagtaa	4380
agcagatggg	cctcaccaat	gtgggtgcac	attgtttaat	ctgttgagt	cctggataga	4440
caaaaaaggc	aaaagaagg	tgaattccct	ttctcgtctt	aagctgggac	agccatcctt	4500
tcctaccctc	agacatgaga	ggtttggtt	cttggtatctt	tggtcccaag	ggctgacact	4560
ggtggccacc	tctggtttca	ggtctttggc	cccagattgt	aagttacacc	atcagcttcc	4620
ttggttcttg	ggccttgaga	ctcaagctaa	aatacactac	cagcttccct	tggttctctag	4680
tgtatgggaca	gcaaatcatg	aaaccttctg	cctccataat	catataagtc	aattcccgta	4740
ataaatctgt	gcttatatat	ctatagcttt	ccttttggtc	tgtttctcca	aagaacctta	4800
atgtacacac	tatatgacct	aacctgtagt	aatgataacc	ttatgcaggt	ttgaataaga	4860
tgatggattt	ctcagtatct	gggaggtatg	ggctagagt	atgaaccacc	gccatgagcc	4920
taggactgag	gagatttctg	aaatgtggaa	tatttgggtg	caaaaccaag	agataatata	4980
gccatgtgga	aaacatgtag	aactatcgta	tgattcagca	acccaaccac	tggaatttta	5040
cccaaaggaa	aggaaaccag	tatatataag	agaatctgca	ctcccatggt	tattgcagca	5100
ttattctcaa	tagctatgat	atggactcaa	ctaggagat	tagatgaatg	gacaaaagaaa	5160
atgtggcata	tgtacacat	gaaatactta	ccagctataa	aaaagaatta	gccaaagcag	5220
tggtgtgtgc	ctgtcatccc	agctccttgg	gaggctgagg	tggaagatc	tcgaggccag	5280
aagtttgaga	ccagcctggg	caaaaataa	agactcggtc	tctaaaacaa	tttaaaaaata	5340
ggccttccct	aaaaaaagaa	taaaatcatg	tcattcacgg	caacatagat	gggactggag	5400
gatattactg	ttaagtgaag	ttagccagga	acagcaagtt	aaacccaca	tattctgatt	5460
catatgcgga	agctaaaaaa	acgttgatct	catagaagta	aaaagttaga	cagaggatgc	5520
tgagactag	aaaaggtagg	gagaaggaa	ggagagggaa	aaatttggtt	acaggtacaa	5580
aaacaaaatt	acagttagtt	agggagaatt	aattccagca	tcctgtagca	ctataggatg	5640
actatagtta	ataataatac	tttaattagt	ctcaaatagc	tagaaggagg	atattgaatg	5700
ttccaacac	acacaaaaaa	atgataatgt	atgagatgat	ggatagtgtt	gttatcctga	5760
tctgatcact	ctacattata	tgtatcaaca	catcactatg	tacccacaaa	atatgtagaa	5820
tttttttttg	tccattttaa	aaagataaca	aatttataaa	taaaataaaa	actaaattag	5880
tgttccatgt	aaacctggat	gaactgggtc	cctcacgtct	gcccatctag	atggctggtc	5940
aaagtttccc	aggctccaca	tcaagttggt	ccactgctca	ctggaacttc	cctagtcagg	6000
ttgggcaaat	agtaattttac	agcaatagt	aatttatcac	tgacatttct	tcagttcccc	6060
tctttggcat	ctgcttcttc	ttttctgtaa	tgtgtttgt	tgaaatgccc	aacattcttt	6120
ttcttcccta	gagctattca	gggtgacctt	tcttttcgca	ttttcccatg	ccacttccat	6180
tatatcaaaa	taaaacagtc	ctgtgtggcc	actgctcatg	accttgtttc	ctgccatgtg	6240
aagataggat	cggtgtctct	ttcttctcct	cctttttttt	cagagacagg	atctctccct	6300
gtcaccagca	ctggagtgc	atggcacagt	cgtagcttgc	tgacgcctcg	aactcctgga	6360
cctcctcagc	ctcctgagta	gctgggacta	cagggtgcaca	ccaccatgcc	ttcctaattct	6420
gatatatata	tatatatata	ttatatatat	aaaatatata	tataaatata	tatatattat	6480
atattatata	tataatatat	atattttata	tataaaatat	atatattata	tatatatatt	6540
atatataaaa	tatatatatt	atatatatata	atatatatta	tagagatggg	gtcttgctct	6600
gtcaccagag	ctgaagatca	gctgctcttt	ctaactctgt	gttagataag	atctgtctcc	6660
caggggataa	aatactacct	ggaataaagg	tatctttaa	ataatcccag	agaagaaaac	6720
atttttatag	tatgacagag	gcagagaaaa	cagagaatat	ttgttaaggc	aggactttca	6780
ccactcccag	tacaatcatc	tgtctgttac	ctgcatacct	tacacgggct	ggcactgctg	6840
gggttacaaa	gtagatgcca	aacttcacaa	tggttagatt	catgtttaaa	aagccattgg	6900
atcaaacctt	tgtgaaagtt	tccagctttt	ttctgttcca	aatatgtgtc	cattataaaa	6960
gaatctcaag	agcataattg	ccaagatagt	ctatgtccat	gagtatttca	acatctctca	7020
tgaaatctgt	tcccatcatt	actcaagata	ttgtatgaac	agtattccac	ataaaactagg	7080
tgctcaataa	tgattgattg	gccaatggag	ggtcattatt	taatgcacta	caatctttta	7140
tgcaaggggc	ccacaggaat	cagtatgatc	ccataggaat	ccttttcttt	tccattgaaa	7200
aagaacagag	tagtggcttg	tattaggttt	cttgtgtgtg	ttgtgaggtg	gaaagatatg	7260
aaaagaaatt	tgatcagagc	ataaatctga	gcccattgga	taggaaagaa	tgagggaata	7320
aggaagaaaa	cacagattat	agacaggaaa	atcaaaccta	ttaaaactga	taattttcga	7380
atactaaaaa	tgtacattca	tttgaacaaa	aagattctat	aaagcaagat	ttctctgttc	7440
ttaccagcac	taccatgccc	aaactacctt	aggaatgaa	tagcagagtc	aaacttaaaa	7500
gcacctgaaa	tttaaaacaa	aaaccaattt	acattttatt	taagaaaagc	aaacagatgg	7560
gcctgctaac	aatgtcaaag	tctogtttac	aaagaaaaaa	acaaatctgg	aacctgaagt	7620
caaacgagtt	caaaataaaa	agcaaaccaa	taaacagaaa	ccaacataaa	cagaagttag	7680

taccatctcc	ctcagcctgt	gaaattctgg	aacttctctt	tctttctcgc	cttcttcttc	7740
tctcacctgg	aagacgagca	gagtgaacac	atcaggggtt	gtcagttccc	cagatggcac	7800
cacattcata	aaccaccgac	tccaggagaa	tgtaggaagc	ttagttaagg	ccaaagtctt	7860
ctttgatct	tcctcatggg	cttcaaggca	aaagaaaaaa	aagtttgctt	gagaatatct	7920
tcatatctat	tagtttgaac	catgcaaaat	tacagttttt	ataggtaaaa	tgagtgcata	7980
ttggcaattt	caaatgatta	accctaatac	attatgcttt	tgggtataga	aatattcaga	8040
tcctaaacat	atgctgttac	atacaaaatc	aggtatattc	ctgcttctat	aattaaagca	8100
aagagaattt	cttttggtca	ctactccttc	tgacatgagg	tatgaaccaa	gttcaggacc	8160
cctaaaggtc	tgggtctggg	tcatttctcc	acctctaact	tgtgccgctt	tcttggtcag	8220
tcattgtgtt	ctgagctgtc	tcataaaaca	tctgctatga	ctttactttc	tcctgatagg	8280
gtggctttcc	atcgttggca	cttcgttggc	cttattggta	tgctttatac	actggttctc	8340
gtttccaaat	tggcattatt	attgttatga	ttcctgtctc	tctccacat	ttcccatctt	8400
ttcctctgat	tcctctcacct	gtacatttct	tacattttct	cctgtgcttc	cttcttccca	8460
tcacattgtc	ccaagtgtgt	cttctttctt	ctccttgtca	catttccttt	gcccgccttc	8520
acatatgcag	agatggctct	tggttttcct	tctgaaatct	catagtttgg	aggtaaactt	8580
gttagcaagg	ccactgagaa	gagaacaaaa	gggaaacata	agagaaacca	agtcactatc	8640
tcctctcattt	cctggtttct	agaagtaaga	cccaaagaac	tcactgtttc	agtgccttca	8700
gctcaggcca	aactagggtg	atcaaaactga	gcttctgagt	gctgatcaaa	acctataaaa	8760
ccaagttagc	agaccattcta	caaattcttca	ctgtttaaata	ccataaagaa	tgaaaaggct	8820
actaaattggt	aagactatat	gtgtgataat	taaattttatg	catcaacctg	gctaggctaa	8880
aggatgacca	ggtagctggt	aaaacattat	tctgggtgtg	tcataagag	tgttttcgga	8940
agagatcagc	atttgaattg	gtgaacttag	taaagcagac	ggctctcacc	aataagggca	9000
ggcatcatcc	aatctgtcga	aagcttgaat	aaaacaaaaa	gaggaaggga	aaatttgctt	9060
cttttcttct	tgatctagta	tatcatcttc	tcctgccctt	ggatgtgagt	gggccttcag	9120
acttaaacca	ggagttacac	ctttggcttc	cctggttctc	agttctttgg	acttggactg	9180
aattacactg	ccaggtttcc	tggttctcca	gcttgcatat	ggcagatcat	gggacttctt	9240
ggcctccata	attgtgtgag	tcaatttcca	ttttatttac	atatccagtt	atgcattgct	9300
taacaatgga	gacaggttct	gagaaatgca	ttgttaagtg	atttcatcat	tgtgcaaaca	9360
tcatagagtg	taactacaca	aacctggaca	gcatagacta	ctacacatct	aggctacatg	9420
gtgtagcttg	taacctcatg	ataagtatgt	ataacatcat	gataagtatg	tatgtatcta	9480
ccatatctaa	atgtagaaaa	ggtacagtaa	aaatatggta	taatcttatg	ggatcaccat	9540
catatatgca	atcctttgta	gactgaaatg	tcattgtgta	gtgcatgact	gtatacgcac	9600
acatacacaa	acacacacaa	atatactatt	ggttcttttt	ctctgaagag	ccctaataca	9660
atatgttata	catttatatt	gactctattt	caaaatttat	ggttttgggtg	aaacatatgt	9720
ggagatgggg	catagggtgtg	tgaactggga	tagtgtcctg	ctgatgaatg	gggtgggaggc	9780
atcatttggg	acaagcccag	ggcatcagct	tatagatatac	aagagctcaa	caagagcact	9840
ttatggcaaa	acctcccaca	agacctctca	gaagttgaga	aactgctaaa	agtttcttta	9900
tgacagatga	catttatgga	taaaataggg	attagcagga	ttcttttaaat	actttcgaac	9960
actaaaccttc	atttctacca	ggcagtgagg	ccccaagtgc	agggccatag	gaagtacaag	10020
tctgggagat	actaggctgc	actgtctgta	gagaatctga	aaaaataata	gagtcactga	10080
aatgcagttt	ggtataatta	ttgccatgca	tcataattct	aatcatact	agtgggtcaaa	10140
tactcttccc	tgaaaaaaca	ttttcttgggt	ttgaattcta	aataattggt	gtgggtcacca	10200
ctgagctttt	aaatatataa	atactttcaa	gtttgcatat	ttttattacc	tgttccttaa	10260
caaacattga	attcaacatg	aaaatgatta	tgggaaacat	tcgggtatac	agtccttgac	10320
tcttaaggac	tcaggtaaat	acttagggta	tttcatggcc	ctagtctttg	gggtaccaca	10380
tgtttcttct	tcaaatcaca	gattcaaaat	caagaatgat	aacacagtga	ttgtgtagac	10440
aaaaataagt	aaccaaattt	gcttgcttct	gtcattctat	ggaaccactg	agagttttta	10500
cttgtgctta	aaattttgaa	tagtaaaaca	gagtgtcaac	ttcatgctgg	aatatttttg	10560
gctttttaga	cacaatttta	agtacatgaa	gtattttttac	aagactaagt	aacatcactg	10620
aaattacagc	tttcttcttt	ttaaaactgg	tatttgttat	aaaactaaag	agcgaatcaa	10680
gaaaagcata	attattactg	attattacag	gattattact	gaaaaagaaa	tgtacggaat	10740
agaggaggaa	ggagtttaac	aatgatccac	tctgggtgtt	gaaaacacca	ataagcctgc	10800
ttccaggaag	tgctaagac	agagctggct	cagcttgctg	ggtcacagca	tgtaaggaaa	10860
ctgctgggct	acatgccacc	atcctcagtt	gtccagatag	ataatcccat	agccccatgg	10920
ggaataatc	tttaattatg	atatagctga	caccattcaa	agcactatgc	taagtccttt	10980
atgtgaatta	acttttgtca	aattttattt	tcataaataa	cccaaataatg	tataaccacta	11040
ttatcctacc	ttaaagagga	gaaactgagc	tcctaaagt	taaatactta	acccaagtta	11100
agactgctag	tcaccctagg	ctattaactc	agtcaggtcta	actcaggtat	aataacatta	11160
tgctactgtt	tgacagcttg	actatgcctg	aagtataacg	tcatgctatc	taactaaaaa	11220
gctaagggaa	ataaaatgag	ccatagggct	caatttcata	aaaggagaga	aaatactggg	11280
gaaaagtgat	aatgcagagt	ttaaaatatt	tttgtaaaag	tgccagagat	tgagtataac	11340

aagtgtgacc	aaaaaaaaa	aaaaaaaaa	aaaaggaaga	aggtaaaaa	aagagggagg	11400
tctgagaaat	agaaatatca	gaggaaggaa	ataaaggagg	gtgagagtaa	attctctttt	11460
agcattcaga	ttccacagat	tcacaaaatc	acatttcttt	ttttaccaac	taaggaaaaa	11520
taacacttga	cctaacattt	cattgcagtt	agctaaagga	tgctagaaaa	actatgttgc	11580
agtggtttgc	tctaatttct	tcaggaatag	agaaaagtga	caaaaagatc	agagaagaga	11640
agaaaggaaa	ctatcagaaa	aatacagaat	tggagtagga	tataacatat	ttgggttgaa	11700
ggtaaaattt	tatatgttaa	tcttaagtat	cttgctactt	cagtttggtc	cctggaacag	11760
cagcatcaga	atctgccgag	ggcttgtaa	aaaggcagaa	tctcagggtc	catcccagac	11820
tcactgaatc	agaatataaa	tactgacaag	atgccccggg	attcatatgc	acagtagagc	11880
tggcgaagtt	ccattgtagc	ctgtgattgt	tttctgcaac	ttagtatttc	tgagttttcc	11940
caaggaagaa	aaccaggcc	ttagcttctg	gcagacttgt	gtttctcctt	tacttactag	12000
ctgcagtact	catgacaa	gaaatcaaac	tttatgtgcc	tgagtttcct	catctataaa	12060
atggagacta	taataatcat	ctcctaggct	tgttttgagg	atgttcaaca	aatgtcctt	12120
tcattcctct	atttacagac	ctgccgcaga	caattctgct	agcagccttt	gtgctattat	12180
ctgttttcta	aacttagtaa	ttgagtgtga	tctggagact	aactctgaaa	taaataagct	12240
gattattttat	ttattttctc	aaaacaacag	aatacgattt	agcaaattac	ttcttaagat	12300
attattttac	atttctatat	tctcctaccc	tgagttgatg	tgtgagcaat	atgtcacttt	12360
cataaagcca	ggatatacatt	atggacaggt	aagtaaaaa	catattattt	attctacgtt	12420
tttgtccaaa	aattttaaat	ttcaactgtt	gcgcgtgtgt	tggtaatgta	aaacaaactc	12480
agtacagtag	tattcagtag	agtattttaag	cccctgtact	taaacatatt	cctcgtacca	12540
atgaagttac	atgaaaagca	aatttgtgtg	agatatcgta	gatggaagta	aattagtctt	12600
tatgttcccc	acaaattgaa	atgcatttca	aaaactctgt	gtgtgtatgt	gtgtgtgtga	12660
cagagtgtgt	gtgagagaga	gacagagaga	tacgctttgg	ttgcctccat	aagctggctg	12720
ctatgattaa	taagaccaag	ttttctaaag	aaaatgagat	cataacaaaa	gccctcttta	12780
tgactatctt	ttatcagggg	caaaaaggaa	agagacaaaa	cagcatgaaa	tgatgagacc	12840
aagtgatgaa	aattcattca	caatgattgc	tttcaagagt	aatttctcct	gggtaattca	12900
gcagcctgtt	actatggctc	tctggagtga	tagctaattgt	aaatgaagcc	tctaaaagtg	12960
gattatcctg	acaagaatat	actcagccaa	taatgcaaca	gaaatccatt	caaagcattc	13020
gggaaaaatt	caaaagaata	aatattcttt	tttttttttt	aaagttaatg	acctacgatc	13080
catttcttcc	ctgactaaca	agcagcaagc	acttaaaaa	atccagccag	gatgaaatag	13140
aaaccacact	gacttgttaa	tatttttgtt	tggtcccagg	gactcagatt	ctaagccaaa	13200
ttctttgaat	gatcttggca	aatgtctcga	attatttttg	ccaacttttc	tttatcttgg	13260
aaaaaaagtt	tcatgaatgg	gtgtcaaaat	tgattagttt	taaaaacctt	tcttgcagat	13320
acgtatggca	ccctaaaact	gtattagaaa	aaaagtaagt	actctgtagt	gtgaaaaatt	13380
cttaaaggac	accctctttt	acaaactcac	aaaaacagcc	tttggaaatac	ccacatgaag	13440
tagctgttgt	tattgctttc	tatataccta	catcttgtct	attataaaaa	gactggtttt	13500
tggcaggtgt	ggtggctcac	acctgtaatt	ccagcacttt	gggaggccaa	ggcgggcgga	13560
tcacctgaga	tcaggagttc	aggaccagcc	tgactcaatat	ggtgaaaacc	agtctttact	13620
gaaaaatacaa	aaatcacccg	ggtgtggtga	cgggcgcctg	tagtcccagc	tactcgggta	13680
gctgaggcag	gagaatcact	tgaactcagg	agtcagaggt	tgcaagtgagc	tgagatcatg	13740
ccactgcact	ccagcctggg	tgacagagca	agactccatc	tcaaaaaaaaa	aaaaaaaaaa	13800
gactggtttt	tcaacagcta	ttcccacccc	tctgcatgga	aatattcacc	cagtcaattg	13860
ttttcctagt	ttgggtaatg	gccctctggg	caggactgga	gtggggcaca	caggagaagc	13920
tgcaaaacta	gtttagaagc	atgtctggga	aatgtcatgc	aagaaaagac	atattttaag	13980
gtaggctttg	catgaatgga	aaaggagagt	aattctatgt	agagcagagc	ctcttacttg	14040
cagtggagaga	agcaaaagtg	gggaagcaag	aggaattatg	cttttcatca	gccaaatttg	14100
caggtaggag	gattggctca	gtcatcttgg	ctgaggctca	tgaaccagg	tgtaaagaaa	14160
gtggactaga	ttaatttcat	ccattacagg	aagaggagcc	gtgaaagata	atccagaaat	14220
cattgggatt	tgatggtaga	aggtattttg	ggactattcc	atttgaaatg	agaaggtacc	14280
tgacattctt	tgaattcctt	tcaagcaaag	gattaaattt	acctatgagt	tgactcagaa	14340
aaaacataaa	aagtattgtt	gctctgctca	gagttttatc	taactcattc	tcacttctta	14400
ttccatgatg	aaatgacata	aatgaggttt	tttattgttg	ttgttgttgt	tttctggaca	14460
caaggcaagg	tagctacctg	ggcagagctg	ttttatttct	ctatgccgtg	gagagaaatt	14520
ggtaatttgg	ccatggaagg	cagtcattaa	gatgttccca	tgcgagtga	ctttccagg	14580
ttcccagctt	ctgcatectt	ccctgtccct	caattccatt	gttggtgatg	acaatgtctc	14640
tcccatcagc	ctcatgaagt	tctctctcat	ttattaaaa	ttgctttcag	gaaaaatttt	14700
gaaaattgtg	ccagtaatgc	ctgattggcc	ccttatccta	aaggcttaaa	ctggaggaag	14760
gaagctaaac	tgagaaatct	tgcaaatcat	tgagccaaaa	acgtattaat	agcaagatct	14820
atcattttatt	gactagtatg	tggcaggcag	tgccctttta	tttaggcagg	gagagttgat	14880
ggggggggcg	gggttcacac	atcttaagaa	ggtgctatct	cctcctatat	aatcatgta	14940
agtcaagaga	gtaaggaatt	gtctttgttt	ggttatattc	aggggattag	agtatacagt	15000

agaagatccc	aagaaacctt	gggatcattt	tagactaaga	aatgccaata	cgcgcggg	15060
cggtggctca	cgcctgtaat	cccagcactt	tgagaggccg	aggtgggagg	atcacagggt	15120
caggagattg	agaccgtcct	ggctaacgtg	gtgaaaccct	gtctctacta	aaaatacaaa	15180
aaattagccg	ggcgtggtgg	cgggcgcctg	tagtcccagc	tactcgggag	gcggaggcag	15240
gagaatggtg	tgaactcagg	aggcggagct	tgagtcagc	cgagattgcc	ccaatgcact	15300
ccagcctggg	cgacagaacg	agactccgtc	tcagaacaaa	acaaaaggaa	atgccaatat	15360
cagcagaaat	agagccaaat	catgaacata	agctaaacaa	atgttggcag	tgtagcctag	15420
tggttaagag	agcagactct	taactagaac	actgcactcc	atgtcctcac	tgtagaccct	15480
cactgtgggg	ttctaattaa	cccctgttac	ttaccagtgg	cagtcttaag	gcattcctta	15540
agttcgttgt	gcccccaattt	gttcactctgt	agaaggggta	ggatgacagt	agtgtttact	15600
ttataggcct	actgtgagca	ttaaatgagt	tactactgta	tttgtaaagt	gcttaaaatg	15660
ctgtctcaaa	agagtttggt	aaacacttaa	gaactgattt	acttgcatct	aaactgacag	15720
ctctcaataa	ctggaaatga	tcaagcatag	gccctggaat	ataagcagggt	ctacatgaag	15780
gcaaaaatgt	tcgtttcttt	tgttcagccc	tgtgcctaga	tcaatatcta	gtgatcatgc	15840
tcaagaaata	ttgttgaatg	aatcaatgaa	cctaccgagg	tagttacata	aaagagttct	15900
gcatgagtac	aaatctgggc	aaagtgcact	ccaaggaaat	ttccactttt	agattctgtg	15960
atttccttaa	ggaactgata	aattgggtgtg	atacaatgta	aaaaaatgtg	cctatatgat	16020
ttgagaaaaa	cttattttct	ctccctcttt	ttctctctct	tccttctctc	ctccctctct	16080
tccttctctc	ctccctctct	tcctctctct	ctccctctct	tccttctctt	tccttctctt	16140
tttcttctct	tccttctctt	tttcttctct	tccttctctt	tttcttctct	ctcttctctt	16200
tccttctctt	cttcttctct	ttgtgccttt	cttcttctct	ttcttctctt	tctgtccttt	16260
cttcttctct	ttcttctctt	ctgcctttct	ttcttcttgt	tttcttctct	ttccttcttt	16320
tttcttctaa	gcagaccatg	tctgttagat	gaatgccttt	ttctagttaa	aagggttaaac	16380
aggaaagtga	agcacaatta	tcaagggtct	ccagtcactc	ccacatgttc	ttaatcatta	16440
tcttcttcta	cagtttcata	tctccaggcc	tttcattggg	tcaggttggc	atttcgctgc	16500
cctttatgtg	tgtgacaagt	gaaaataaag	aaagaaaaaa	actcaagtga	agaaaaatcag	16560
aatctgcgca	gcagttcctg	ggcgtttcag	ctgcttccca	catcacctgc	ctcatcaagc	16620
cccagcatcc	atctccttgc	tcactcttaca	ccctgtgtgc	atgacaggcc	caccattcat	16680
ttatcagagc	aaaggctctc	ccactattct	gggtcacccc	cctacttagc	cagatatata	16740
agaatatctg	cacggatgac	ctgcctcacc	tgggagctca	gaggagctca	gattccatta	16800
ctatcgcacc	aaggacagat	ctcccagcaa	gaatgacaga	aaagactaac	tgcccccaaa	16860
atctcccttc	caaaacacag	ttctcttaat	tctcccaaga	aaccagaatg	tgactgtcta	16920
cctctctaag	gacctgaaaa	caactggcca	tttcagctat	ttaaatcaac	tttaaaaaat	16980
ccaaccgcca	aaatattaaa	ccatttttgt	tggaaatgata	acataactaa	cctgctgaca	17040
gctgcttctg	ctaggtgcaa	aaatggaaaa	aaaaatactt	ctaatacagg	caaatacctc	17100
tacctttggg	attctaaatt	tactcatatt	ctcaaagaaa	tatatctagt	catagtgggg	17160
aaaataggat	tattccttta	gctcgataag	caaccagaag	ttcttctctc	aaatcttgac	17220
atgtaataca	ctgaaatttg	atttttgtaa	aactgtttcc	tatgaagcta	tctctgcctg	17280
aaggattttt	tttttacaat	ccagactata	gaaggaaatt	cacaacctgg	actttcacct	17340
ccattgggtca	gagttttact	gaccaattcc	cacctctgcc	ttacacctaa	cggaggttta	17400
tgctgttttt	ctcttcacat	acccaacag	ttacaaatgg	ttgttattat	taagcatctt	17460
ttattttgtg	gcctctgatt	acatgggtcc	ctaaattttg	acctaatac	aaaagattgg	17520
taaaatttct	taacatatta	ataatatttt	gtttatgtgt	caatatctta	gcatgtatca	17580
attaagacag	aggtcttaac	gttctctttt	tgaagagaaa	tattaggatt	cagagatatt	17640
aagagattct	cccaggatca	cagttaggtg	acagagctgg	attttagtcc	aggtctgtct	17700
acagctctaa	cgtatataca	ccctttgtat	aacatgtcac	gaattcagca	taaagggatc	17760
ttcagtgatc	taagtcaggg	gtcagcaacc	ttttctaaaa	aggaccaa	agtaatat	17820
caggctttgt	ggaccctatg	gtctctatca	taactgttca	aatcaccatg	tagtgtaaaa	17880
ggagccataa	gcaaaatata	aactaacgaa	tgtggctgtt	ttatgggatt	tttttttaac	17940
tctttattta	caaaagcagg	tggcagatca	gaactcactt	atgggccata	gttctctgac	18000
ccctgacctg	agaaaatctt	atattttatg	acaacattta	gactgtgact	tgccaagtta	18060
gaacaagaag	ctctgtcaac	tgaaggtcaa	ggctggagtt	ctgaaaagcaa	agagctgtct	18120
ggtgttaatg	ataagtga	tagttaaagt	tagaagatcc	cagttataag	aagcacaaag	18180
aataatgacc	atagactcct	gaacaagaat	gtctggactt	ctggcttagg	cactcttggt	18240
gtatggtcca	ggccaagtta	cctaattctc	ccaggcctcc	attttcttat	cattaaatga	18300
agataataaa	agtattttcc	tcagagagct	gtaagaataa	actgagctaa	cccatgtcaa	18360
gcacatagaa	tagggcccag	cctatatata	tttatcaata	aatgccagct	acatatagtg	18420
tctctatatt	tttattcatt	atcataaaat	gtttatctac	agattggcat	tgtaaggatg	18480
gagttaaaat	tgtatgtatg	tgaaggga	ttattctctg	tactattgat	ctgcatcaca	18540
ttaccccaaa	tttgatggct	taaagtaaca	acattcattt	tgcaaacaaa	tttgaaattt	18600
gaggagggct	tgtctgggaa	gacttgtctc	tgccctatgt	ggtatcagca	gggggaggct	18660

tgacggactg	gcacatgccc	ttccagaatg	gcccactcgc	atgcctgcc	agttgggtgct	18720
ggctcttgge	tgggagctca	gctggggctg	agtgcaggag	tccttgggag	gttccttgtg	18780
gcctgaactt	cctcaccaca	aggcggtcgc	gggtcgagag	tgagcatttc	aagatagagc	18840
caagatgaca	ctgtattact	gtgtaagacc	cagcctggga	attaatgtag	cctcacttcc	18900
atcccactct	atttttaaaa	agtgaattat	taaggtcacc	ccatattcaa	gggtagatga	18960
attagacttc	atctgtatta	agaaaaatgt	ttttaaaaa	tgtagacatg	ttttaaaatt	19020
ctaaagtcca	cttactggct	gcagattatt	tatatataca	tgcaagatac	actcctacat	19080
tctcttctta	gaaggctcag	ttgcaggtag	agatgaagct	cttcaagtga	gatttcttat	19140
gtatttatcc	tctcaatctg	aagacttgta	aactaagaga	caagttatct	gcaacctaca	19200
tacgcaatat	tcaatggtaa	agtatacata	ggacagccac	tacagacact	cttgttttaa	19260
atagaggaaa	atgagagcac	ataacagtca	tggctcata	gcaactctga	tatccagaca	19320
gcaaacacaa	gcaggctctt	tttttaggtc	gctcctact	gcctggattc	cctactgtct	19380
ttgggtcttc	ccctcagggt	cttgggtctt	ggacctctt	tcattttaata	ctatttctgt	19440
tcctttaagt	tcaagctggc	aaaatatgat	tgtacaattc	tgtttaaaat	tccaggactt	19500
cctgtgatcc	ttattgggga	atactccatt	agacaagaat	ctctttgaca	taagccattc	19560
tctacctgag	atccctgtaa	ggctgtgatg	ggaccacata	accttaaaat	tattagaaga	19620
ctcattgttt	actgagagaa	tatgcctagc	atatgcttag	atccttagag	gaactctgtt	19680
tcaaagggtc	tatgagacat	taccttatat	tttcttaagg	tacaaacaaa	aggtctttgg	19740
cttttgagtt	tgatcttga	gctgacacct	tttcttaatt	tgagaatccc	ctgctctatg	19800
gagagactga	caaagagaaa	tagttttata	tttgaatgta	acatcttggg	tctttaatag	19860
attatcttaa	aattttcctg	aaaatgtaac	agttcctttt	tttaaaattc	attctcccta	19920
cacactttat	atatatgact	aaaagaaact	ccctggcatt	ttcaacattc	tggttagaat	19980
ttttcttagc	ccaatctacc	agttttattag	gtatatattgt	attttccacg	tcactgtagg	20040
tgacaatctt	gctaaacttt	ctgccactac	ataacaagga	tcctttcttc	cagttttcag	20100
taacaatcta	ccaatagcct	cctcaaggct	taccaagctt	ctactaaaaa	tctctttcag	20160
ccccttccaa	cttctgcccc	ccaactagtc	ccagaagcaa	tacccatggt	ttaggtttca	20220
ttatagcagc	atctagtctt	gagttctcag	aacctgtttc	gatcatgttt	caactgtttc	20280
aaaacatccc	aaaacagtag	cttaagacag	taacacttat	tttgctcaca	ggctgtacgg	20340
gaagcatggc	tggagagggt	tcaggaaata	tacaatcatg	gtggaaggca	aagagggaag	20400
aggtatgtct	tacatggcca	gagcaggaga	aagagaaggg	gaagtgcca	cacacttttt	20460
agaaacccag	atctcatgag	aacttactca	ctatcatgag	aacagcaaag	aagaaatctg	20520
tctccatgat	ccaatcacct	cccaccaggc	ccctcctccc	acactgggga	ttacaattcg	20580
acatgagatt	tgggcaggga	cacaaatcca	aaccataaca	gttggaacc	ctttttttaa	20640
gaaagtaatg	acatcaactc	cttggggatg	tggattgggg	gagaatattg	gagaggatcc	20700
aggggaagtg	aagatatcaa	gttctttata	cataaataga	tctatctttt	tagggaagta	20760
gaatatgtca	tttaggatag	ggaaagtgtg	agatgttact	ctattcagct	ttagggaaac	20820
tccaagatgt	aacactatgt	cctgtgaata	ttagtcttgt	ggaaagtgtc	ccattggaag	20880
aagacagaaa	atgtcctgct	gcacagagat	ctatatcttc	atgccacgtc	ttgcaaatgc	20940
agtcagggag	cttgctctaa	attcgtgttg	ccttttagga	agttacagag	catatttctt	21000
tatatcttct	ttatttttgc	ccttgatagg	agactgccag	ggattcagtt	attataattt	21060
tacattttat	tatccttttc	tgggagcatc	tcataccttt	tttactagag	aggccaaggt	21120
cacaattaag	attagggaga	caatacagag	tcaagacacc	tgaattttgc	cacactactc	21180
ttggggccact	ttaattgggtg	tttctgggcc	tcatttttat	tactattatt	attaattatt	21240
ttatggtgtg	tataggagct	ggaagatagt	gacttctatt	agaatttatt	cttataaaca	21300
atgctgtatg	gctccttgat	taagaagtaa	atagaaaaca	aaaagggtct	ttgaactgct	21360
gtctcttcta	ggtcttctca	gagtgctact	gggaattaga	tgggattaga	tgggtgaatt	21420
tcctttttct	tttgtttgag	aggtctagat	ttttgaagta	tctaaacat	cttttagatct	21480
aagttaatct	ccaaaaagct	ttattgcctg	aaacattcca	gttctaggaa	gtttgcattt	21540
ctctgagtgt	aagggctctg	tctgtaactc	cttgagatca	ggaatgtgtc	tttttatccc	21600
catagttcag	cgctaccct	agtgtcttcc	ctacaagtga	tcattgaatg	aagaaaggaa	21660
taaatctctt	cttctctaac	atacatcaga	gcctctgac	ccttattggg	tcccccttcc	21720
ctattgatat	acctctccat	ctccagacat	ccacatggcc	tcctaggggt	atgggattct	21780
tactcttct	cctggtacaa	gctccttgcc	tgaatttttt	ctgacccacc	cccatttcta	21840
accgtcacag	gtttagattt	tctgaagcaa	acatacagac	agagttcgag	ttacaagaca	21900
tttaatatga	atcgactctt	gtgaaaggaa	gaggggaaga	tgtgggactg	ggctcagtga	21960
gaagccaagc	tgtgacacaa	tgaaaacctc	agccaccct	gcacagagct	caggaatgag	22020
tgtcacgtgt	tagagctgac	ctgggtcagg	tcaacattgc	tgggtgtttc	tgcttctgct	22080
tcgttcagtt	gtcagatgca	agctgacctg	ggaagggttt	gactgggcaa	gatggctgtc	22140
tgagatgag	gtcaatcaga	aggggtgac	agccaaaagt	ggtttgctca	tctcactacc	22200
cacagctggg	tggcaagtcc	ttcaagaggg	agctggatga	tgcttctcca	cctctaccac	22260
accaaacaaa	gcatttcttt	cttattcttt	ttcttctgt	ctcccttccc	cctccccctt	22320

ctccccctttt	ctcctcctcc	tectcctcct	ccttcttctt	cttcttcctc	ttttttctcc	22380
ttcctctttc	tttagagaca	gggtcttgct	tggttgccca	ggctggagtg	cagtggcaag	22440
gtcatagctc	actgcagcct	caacctcctg	tgtttaagca	atcctctcac	cttggcctcc	22500
caagtagctg	ggactacagg	catgtgctat	catgcccagc	taatttttta	atttttacta	22560
gagacagtct	cactatgttg	cctgggctgg	agcatttcta	ttcatgggtt	ttgttcctct	22620
ggagctgtgc	tgttctcttt	gataacatga	atatggattt	gtatttacct	ccttctggcc	22680
atcttaacta	atggcccaaa	ggaaggccag	tggtgaagat	gtttctctac	aggaagttt	22740
ctaagtgttt	ctcagcctac	atgtgtgagt	gtattaatgg	gaatactggg	gataattgga	22800
tagaatgtat	gaaattagga	ttggcctaaa	ctatctggtc	caaagtttat	cctagtggcc	22860
ctaataacat	ctccagtgat	ggtctttcct	gaggttatag	aatgagctat	aggaaagagt	22920
gatagtacag	gaatgagagg	attggatcgt	cttttctctt	tgttctattc	ctctatgatt	22980
attttgatttc	ttttttttgt	gatggagtct	cactctgtcg	attatttgat	ttttgtttca	23040
gattaaggcc	agctttcttt	gactgagcag	acccacgtat	atgggaaacg	gccacacacc	23100
aaattaaacc	attcactgac	acccgggcac	acccattttc	accactctgc	tcttgcccgt	23160
gccattgcct	ttgcatccaa	gggattcttc	cttctgtggc	aatctgtatg	acatccttac	23220
taagaaggcc	tccctcagtg	cccttgcccc	tctcacagcc	ctttccatgg	ccttctccac	23280
tagactgtga	ggaacccagt	gtctggaaca	gttcatggca	ttgttgtag	tggtgtcatt	23340
ttgtttgtaa	gccttctgct	gtcagtggtc	attccatag	tttgcatatt	tacgcccctg	23400
ctcttttttg	tatgcagtgt	ggaccctatt	gcacggctac	ttgtgtttca	gtggctttgt	23460
cctattccac	cttctatgcc	acttactctg	aatgcctggc	acattggtag	tgccccaata	23520
ccaacttgga	aagtgaatga	atttgttttc	tcattacctt	ttgaaatgcc	tatacagatg	23580
ttcctccact	tatgatggag	ttcattctga	aaagctcatg	ataagtagaa	gatattatgt	23640
cagaaatgta	ttttattgta	ttttttatct	tttaatcttt	ctttctcttt	ctttctttct	23700
ttcattcatt	cattctttct	ttctttcttt	ttctttcttt	cctttcatct	ttcttttagt	23760
agagacaatg	tctcactata	ttgectaggc	ttgtcttgaa	ctcttgggct	aaaagcaatt	23820
ctcctgcctt	gctctgccaa	agggctgaga	ttacaggcat	caactactgc	acctagttag	23880
aaaatgcatt	taatgctcca	ataaaccat	cataaagtca	aaccatcagc	agtcagtgc	23940
tatctgtata	ccttaggact	aagttagaag	ttggagatgg	ttaaagttct	aaattagtac	24000
aagcacagga	gctcttattc	ctgccatttt	atgaaagtca	tgtgagtatt	ggtgaccatc	24060
actaatgggg	ctggtagaaa	cttttaggacc	agtctgtctt	agtcagctgg	gctgctacaa	24120
taaaaaacca	tggactgagt	gggttaaaaca	atggacattt	acttctcaca	gttctggaag	24180
atgggaagtc	caagatcaag	aagccaacat	ggctggattc	tgatgagggc	tttttctgga	24240
atgcagatga	tcaccttctt	gccttgtcct	catgtggaag	agagagcaaa	ctctctagtc	24300
tctctttctt	ttcttaaagg	acaataatcc	cattatggag	gcctcatgct	catgacctca	24360
tctaaacctt	gttacttctt	aaaggctcca	actccaaata	ccatcacact	gaaggttagg	24420
gtttccacat	ataaacttcg	gagggacaca	aacattcggg	tcataacaag	tgcatcttac	24480
aggactccag	ggagtgtgtt	ggctaccctt	catcccacac	agctgcaaca	atcctaattc	24540
actgatatac	agtcaccata	tatttgaaat	tactcagtt	cccaagaaaa	tatctgcttt	24600
cataatcatt	tctcagctaa	ggaatgaaag	ctatgataag	aaagtctagc	acttagtggg	24660
aaacaaactt	gaattaggcc	aggaccctt	aaatattgtc	actatgaaat	ccaagggcag	24720
tatttacatg	cttttccttc	atgatgacct	tagatttctc	ttctttatgc	tgcaatcaac	24780
tgaaaacatt	ccactgttgg	taatccttta	tcttctaaca	aatctaattt	agtagatctt	24840
gttactcctt	atttttgaaa	atattaaact	aaacaataag	tgatctttca	gggaaattct	24900
gcaaggaag	tggcttctag	agaatgcctc	ttctccaaaa	ggcctagttc	tatttctgat	24960
atctaatac	ttctgtctcc	tttaaataat	tactgaaact	agtttattgt	ttaaatgacc	25020
tcatttgaaa	gtgggacttt	cccctttgcc	ttcttccctt	tttaatttctg	tttttaatta	25080
atgtaagtgc	agtcagaact	gacgggtgtc	acctttctctg	taactgtgat	tgccacttgt	25140
ggacattctg	tactcataag	agtgattctc	acaagttctt	taagtttttt	ggaagtgc	25200
acccaacagg	aagaaagaat	aggactccct	gcattctatt	ttgctctgca	gtgaagatgc	25260
attgttgagt	aatttgcttt	acattcagat	ataccagat	tcccatcttc	ccaatccatg	25320
tgactttggc	aagttacaac	tctttataag	cctcaatgcc	tgcaactgaa	taagaaagt	25380
gtaaccatct	catggggcat	tggtgaaatg	agataatact	tatgaagtat	gcagcacatc	25440
gccaagagcg	taacgtgttc	aatcaatgat	atccattatt	attacatgtc	aaatcacatc	25500
tgacttctgt	gattctactg	gggaaaagga	tgcaaaatcc	tcttgccgtt	ggggtgtcaa	25560
atattttagt	acttatttcc	aaattcattg	gaggggtctt	tccctgaagt	ttacagatgt	25620
ggaatacacc	attcaggcag	gctttcttca	gtagtgcagg	tcatttcagg	ggttctccac	25680
ttctataaat	atttcttttt	tttaccact	acagacacat	ttactggcaa	ggcccatgat	25740
aacaaaaatga	aattaatata	ggctttacag	tcagagacct	aggcgacctt	gggtaattca	25800
ttaatatctc	tgagcctcta	tttctctcag	tttgaaagag	gcataagaat	tgctctgaag	25860
aataatcatg	agtataaaat	aagataaatg	agataactat	gtaatatgca	ttagttagc	25920
accaagccat	atacaaaagg	acaataaaga	gaatctaata	tgtttctttt	attccactg	25980

ctatgaactg	aatgtgtgtg	tccccgccc	accgcaaatt	catatgatga	agccctaate	26040
cacattgtaa	aggtatttgg	aggtggggca	tttgtgagg	aactaggttt	aaatgaactc	26100
atgaagggtg	agcccccatg	ctgagatcaa	tgccttatt	taacaaaaat	cagaaacaac	26160
aacaacacaa	aggaagacag	catgaaatct	ctctctcttt	ctcttttccc	ctctccctct	26220
gccttcatcc	ccatcgccat	gcaccgcaa	aagaaaggcc	atgtgaggac	ataaccagag	26280
ccctcccaag	accctgaccc	tgctggcgct	ctgatctcag	acttccagcc	tccaggacca	26340
tgaaaaataa	atgttggttg	ttcaagccac	ccagtctatg	gtatttcagt	gtagcagttt	26400
gaactgagat	acctattaac	agatgttcct	ttccttcttc	tccctcattt	ttatggcaca	26460
ttaccagtgt	ttggtactaa	atactagctg	ttaattctcc	cttagtagaa	gacaatagag	26520
taatgctgag	ttggtctcag	ctgtgggaaa	ccttttgccc	agagagaaaa	attaagtctc	26580
ctaacagaaa	gctggcactt	tcccttactg	ggcatcagcg	aaaacctgat	tcccatgggt	26640
aaagttagta	attgctaaa	ggaaacagat	catgtgcatt	tacatagtca	gagatgtttt	26700
atgggattag	gtgggtgatg	aagagtaatc	attataataa	ttttacaacc	cagacataaa	26760
agagaaagag	cagagaaaga	ccagtctctc	tgtgaaatat	ttgccattta	atatcttttt	26820
atctttgggc	aagccaactt	attttcttac	aaataagcag	taactagaca	cagatattta	26880
agaggagaaa	tggaccttag	atgttctgtg	gcataatttt	cctttcacat	aggggaaaac	26940
caggaccaga	gcagcggatg	aagatgccaa	gatcaggcac	tgccaattgt	tttgactacc	27000
tctagggaga	tctctctcatg	actcacacat	tgcctctctc	ttgaaaatga	gtaaaagagt	27060
gccccaccac	ccatctgaaa	taaggtcagt	tataaatgcc	tatagcagct	accttctttt	27120
ttattattta	gaaaatgatc	attggaactt	gttatcagaa	gacacaggct	gagagggaag	27180
ggatgtattt	tttcacatgt	caaggaacca	ggggtaatat	aacattcaca	ttctagagaa	27240
aggtcaaata	ataaattgcc	ataaaataca	gactgctcat	gagctgtttt	gaaatcttgg	27300
cttcttgtca	aaaaactaaa	ggtggaggag	agtaaaatcc	ataactgtgt	gctttttgtg	27360
taggtgtgtt	ggactgaatg	tacatatgat	tggatcatag	agatgatcaa	agaagtgtat	27420
actattcctc	aacagaagag	gacaaaaaac	gaaattgaca	tttgttgaca	tttccttggc	27480
cttagaagtg	tttgtaaatg	acagagttag	gatttgaacc	cagttccagc	gagtttaaa	27540
cccagggttat	ttcccaataa	gccaaagaaa	tttccaatgc	caacctattc	ctggcatctt	27600
tagcagcaag	agtgattata	gtttccacag	acagaaattg	tgtacagagt	tgggtgaaag	27660
aacagtagct	ccaacatgct	tggcatcaag	gctgtgactg	cacgaactat	tccaggaaaa	27720
ctagtcatca	caggtgattg	aaaagagaag	atatctttta	gcagactaaa	atctgcaggc	27780
atcttcatca	ctccttaatg	cagacagaca	aaactgattg	gctagactaa	taagatcaga	27840
ctttgttttg	gtctattaag	acacgcgaga	agacaagatt	tcctaaacat	27900	
agactatgaa	tagactatga	atagtcaaga	gggagggtcg	ggaaaaggga	attcccctta	27960
acatggcaga	tcaaataattt	aaattttaat	gacatattag	atggaatata	agaagttttt	28020
gtgttatgaa	ataaaaagaa	ctcattcttt	taaaaaatgg	caatccaata	tttatacacg	28080
ttcttctatc	ttatagcatg	actgaattcg	aagtgttctt	gttcattgtc	ccttctctga	28140
gaaagagtaa	gtgcctccac	aaacacattc	actgggagtc	cagaagtggg	gagtatcgtg	28200
ttttcactgt	gtagtgttgc	ctggcccaaa	caagtaggag	ctttctaaaa	28260	
ttaaaggcaa	ctatccataa	aggttccttc	atctgctgtc	tcctagtcca	ttggcaggac	28320
ttggcaggga	aagccctctt	cttgatagag	atattggacc	tccaacctat	ctcttcatct	28380
tctccatac	atctcctcca	tacagggttg	tcaggtcagt	aatccctact	ggccatttag	28440
aagggtgacat	ttatcaatta	taattaggtg	acttaggagg	cttactcaac	tccccagaat	28500
ctcagtgtcc	tcatctgtga	aactggaaat	aataaagcct	atttcataaa	cctaattctt	28560
catattaaat	gggataacat	atataaagca	cttagcacat	gctatgaccc	cagaagatac	28620
agattcacag	agcagatgga	gcactgctgt	gtctcatgat	actcaagctt	gaatatcaag	28680
cttgactccc	attctgtctaa	ttctccctct	ggctactagg	acatcttgac	tttcagggtc	28740
tggactctgg	gocctaaagg	aaagctaaga	ctcgctcttg	tatttctact	ttatgtaggt	28800
agaagctgat	ctgcattagg	gctgggttgc	cccttctccc	catctcaagc	cccagaatct	28860
cttcttgtat	taatctctct	cccaaagttc	ttgctgaggt	cagatttcat	agtcctgctt	28920
tttcagcacc	tgttcttaat	tagttcaaac	ttttgcagg	tttctttcag	gccatatttt	28980
atttgtttcc	tcagaaactc	ctcttagact	acatccctta	tctccagaca	cctgggatcc	29040
acattgattc	ttttattttc	cacaattttc	aaccctgaac	agctgcattg	atgtaagcaa	29100
gcttctaagc	atagatgggt	gggtgggttg	gatgtgtaca	cgttattcat	tcattcagtc	29160
atacatttat	tatttataaa	ctacttattg	aacacttcct	agtgttgga	atttttttaa	29220
aaatctcaag	cgatatgcat	atcaatacaa	aataatttca	gttctcaaga	aacttaatct	29280
tgggaaggat	aaacaaagac	ttataataca	atatggtgca	agatttaca	taaagcagtg	29340
gttctcagca	ttggctgaat	aatagaacca	cctggctata	ttatatagtg	atgcttgact	29400
ccaactcaga	ccaatgaaat	caaaatccct	ctggagtaaga	cctgggcatt	gctatgtttt	29460
gatagctttc	caggtgattc	cattctgcag	ggaggtttga	gaatcccttg	gatagaggga	29520
agtatgggct	gccttgggac	cacaattaag	aggtccctac	tctaattggg	aagggtgctaa	29580
aggctgcaaa	gggaagatga	catctgagct	aagtcttttt	tttttttttt	tttttttttt	29640

tttgagacag	agtctcgctc	tgtcacccag	gctgaagtgt	agtggcataa	tctcggctca	29700
ctgcaacctc	cgcctcctgg	gttcaagcaa	ttctcctgcc	tcagctgccc	cagtagctgg	29760
gataacatgc	acccgccact	atgcttggct	aatgttttgt	attttagta	gagatggggg	29820
ttcaccatgt	tgaccaggct	ggctcaaaact	cctgacctca	agtaatctac	ccaccttggc	29880
ctcccaaagt	gctgggatta	caggcatgag	ccaccacacc	cggccctgag	ctaagtcttg	29940
acaaatgagt	tgcccaagtg	aagatgtggg	aaagccattt	caagggcaga	ggaacaagat	30000
gagttcatag	aactgcctgt	aattttgggt	acagtttaga	gagtgctaaa	tgctaaagca	30060
tttagcctgt	ctctggcagg	caaaaggaga	gtcattgaag	gattgttaaa	aaggaaaata	30120
tatggtcagt	attatataca	tctgaagtac	caagacatca	gttgggggagc	aaattttact	30180
ttcattgcac	cttgccccag	cagaactatc	tccctttaat	aagaccttta	aaactgttgg	30240
caacctccctg	gcttcaaaaca	ttagaaacag	gaaactctac	ccatgagcta	accttcaccc	30300
catcctttcc	tgaacaaaag	aacatctgtt	ctccacttg	gatgtgcctg	gagtaaagag	30360
gggcgcctaa	gtattctcag	ttgactacag	agaaatgatt	ggcgtcacgt	agattagtca	30420
tacatacatg	cagatgatga	catagtcaaa	tgtattgaga	catttttagtt	ccaaaaagta	30480
agtgttttgt	ttctgatata	gtgctgtcac	ccttggccag	ggaaacccag	ggatcatagc	30540
tgctacagat	aaattatgat	ttattccttg	gaagaggggc	tgaactcaaa	agaaattaa	30600
tagtaataac	acttgaaaaa	tctctttctc	caatttcagt	atgcaaatct	atttacata	30660
caaccaagcc	tttgggacag	gttagttatt	gtttgtggtg	gtagtgggtg	tgttgtttat	30720
gtttgagttt	tgtgtttttc	tttaattctg	agatacatta	tcaatttaat	gacaactttg	30780
caggataaaa	aaggaggagg	agaaagaagg	ggaggaagag	aaaggggaag	aggaggagaa	30840
gaagggaaaa	aagagaaaaa	atgattaaat	ttaattttgtg	ttctaattgt	gaaacaatct	30900
aatttcaaaa	atgttaaaat	ctgtgtgttg	ggtagagat	gggaaaagta	tgctgtataa	30960
taaaaaaaaa	atccccagga	ggaaaatgaa	atatttttag	aggtagtagg	aattttatag	31020
tcaatgggat	ccaaagctgt	atttttagac	agagaaatct	gaggcacagg	tagacgagaa	31080
attttcttca	aattctcgag	ggaattactt	agtagtgaag	ctggaatcag	aaacctcatt	31140
tctcagcctc	aagtttggtt	tgttttccat	gacaccctag	aataattcaa	ctaatatctg	31200
agaagcaaaa	aagcatctca	tcctttctac	tcataatttct	ggctgaggga	actattcaaa	31260
gagataaagt	acatgaataa	ataaaaaatg	tgtatcacag	ctaagtacat	cttagcaaa	31320
aaataagcaa	aagactatgg	tgcctcagtt	ctcaacactc	tcttttggtc	atttatttag	31380
tggtgtttat	cttgagttag	tgattcttgc	atatcaaatt	gcattttggt	gtgtcaaatg	31440
gactgggtttt	acttgtttgt	tggttcctaa	agctaattat	tatttttaac	accttttgat	31500
atgctaaatc	aggctaataa	gactaaaaatg	ggattttctca	tgggttttatt	agtcataaac	31560
accatcttac	atatgtgtat	atatgtttta	atgctttcaa	agtgtaaata	aatgattata	31620
cttgagactc	atacaaatat	cataaagcag	ctttttcctt	agtattcctc	ctagaaaata	31680
gatatttcat	tggcagagtc	gtggtgaaag	atgtggtgtc	atacaataaa	tttacttcag	31740
ccagctgtgc	gagcttaagt	tacccagttc	ttttctgcct	cagtggcctc	ttctgtaata	31800
tgactgata	gggactctca	catgcctgtg	tttagaatta	agtcggataa	ttcatgtaaa	31860
gtgcttagca	agtcacctgg	cctattataa	ctagccatta	ttgtttgcta	gaacatttat	31920
ctgttcacaa	attatgcatt	gattgcctat	aataactaaa	cttgagggaat	gcttattacg	31980
taccacgtac	tattcgaagc	acatcacaa	tagtaactca	tttcattttt	ataatgactc	32040
tttgaggtaa	gaactattat	taccatctct	gctttacaaa	tgaggaaatt	gaggcataga	32100
gacgtaaagt	cacttgttca	aagcctagca	tctgggaaac	agtagagggtg	ggactcaaac	32160
tcagaaagcc	tggctatagg	ctctgctctg	ctgtctcccg	ggattggcca	catgctaagc	32220
attggggcaa	cagcaatgag	gcccataagg	tttctgctct	caaggagatc	actataaatc	32280
agccccggga	gatgatccct	aaaagagaag	tacgtcgttt	gcaaaatgac	agcacagacg	32340
cacgagttct	gaactcttcc	tcagagtttg	gggaagttgt	ttagccctgc	tttctgccct	32400
ctctatgggt	cacacacata	ctccgagtga	agaatgacgg	agttacccat	cttctggctc	32460
cttcccggtg	cactggctac	cctatgctaa	aaaccagccc	ccagattcca	catccctgaa	32520
tgtcttattt	acaggttctg	tcattttcat	gataaccttt	ctaattgtct	aaatttcctt	32580
tctgtgggtt	gtcaagcaaa	ttcattctcg	gcagtcagta	agaactcctt	gttcaaaatg	32640
ccacacatta	cagagtcaat	gattctaaac	acattttttt	ttagtgcata	caacttataa	32700
gctgcagtgc	tagaaacaaa	ttcacactcc	ctttacacta	cagtgccatc	ttttccgagg	32760
gtggggctcg	cgtttcttca	ccatccagtt	tctcatatct	agattatgat	ttctgctttt	32820
gccaaacagg	acattgtttg	accaagtcct	tatgggagga	agcacagccc	agtgagcaga	32880
acaatgattt	gaaagtcctg	ggttctatct	tcagctcctc	tgttgactag	ctgcgcattc	32940
agccctagga	aggtcaaaag	cagaatggcc	ttgctttgac	gactttctct	gcattgtacg	33000
ctctagatgg	aaaaatatc	aggcatctgg	ctctcagcct	tgaggagctg	gtctttggag	33060
aatacatgtg	tagtcacact	ttgttttcat	ttttaaagca	aaagaacaaa	acaaactctc	33120
tagggggaaa	atgggctggg	agagaatgga	agagagaaac	ctaataacat	tttaaacaat	33180
ccaggagtta	agggattgct	tgagtcattc	agatcaaaaga	aggctctttg	gcacaagttg	33240
gagctctccg	catttgaaaa	tgggaaggagc	aaagaggagg	aagcttttag	agactttttc	33300

aaaccgggac	ggtagcagat	actcagaact	ctgcctgtag	ggaccaggcg	gggaactcag	33360
gcaggttaag	gcctgtgggc	aaactggaaa	gcccaggcca	cctgtgaagt	ggtctgctgc	33420
tagtgatacc	agatgcttgt	tgtggaagaa	ggaagttact	cttgccagaa	cagctgaatt	33480
ttcaagagaa	actgaacatt	gagatttttg	tcattagaac	cctgattttt	tttacctgat	33540
gaaatctaata	catcattttt	aaacttttta	aggggcaactg	agcaggcaag	atcaaacata	33600
tagcctaagc	aatgtactcc	ccatgtctct	tctatatgta	gatggggtgt	ggttccttaa	33660
gggctttagt	tgtatgttct	gtgaatgaat	agaggctaga	ttccagacaa	aatgatttag	33720
cagtgcactca	acttttagctg	cactttggaa	tacctgggtg	ttacctggga	ctgcagcctg	33780
agcactggga	ttttataaat	gcttctcaag	tgattccagc	atgcaaccat	gaataagaac	33840
cactatgtta	ggatatttgcg	agaaggttta	agcattacag	aaagttaggt	tttaacttga	33900
taagatacat	aatcttatca	agacacatgt	gatctctgaa	catggatcag	ggaattgaca	33960
ttcaatataa	atggcatttt	cttttctttt	tttagactct	tttgagagtt	ccagttaaaa	34020
actctctagc	atcattgctg	atggccgtaa	ttaatatctt	tgggttcatt	agcaatcagc	34080
taacaacaga	catccattgg	caaaacaagg	gtatagtgg	agaggcaca	cacaacgaaa	34140
gccaaaggac	aagtgcagc	cctgagcaca	gataagtatt	tacagggaaa	attaactttt	34200
ttctctgatg	ccatgaagca	aaatagacag	tagagcaaa	ggagcctttt	acaagagcag	34260
ttcaatcttt	gtccacattc	taatggctat	aaaactcacc	atgaagctag	caaatggtgc	34320
atacaactaa	tactttcagt	cttttcttcc	ttatttcaca	ggtcttatgc	attataaata	34380
acacgactaa	gtctcagcgg	ttgtatccat	ggctagtctt	tatttcagct	ttggttcctc	34440
ttgctattta	ctatgtttct	cataaaagat	gttagtgtgg	attgagtata	tgaaatcatg	34500
tacaatggaa	tgttaaattg	ggcattccct	cttgtctagg	tatatcagaa	gaggcccaat	34560
acacttggtg	ccaaggaagc	tcaccttcta	aaaccagatg	gtagcttgcc	tataaaagga	34620
tacactaaga	ggttctaccc	aagttatcac	attggcagcc	tccagcttat	caagagagac	34680
tttaagaac	ttggcagtg	ggatctttct	agagggcggt	gtcaatttag	gaagcatctt	34740
cgtcttcatg	cctcaccttt	ccgataagtt	tgtttcttcc	ttaatgcaaa	caaaatgaga	34800
gacactttga	tgggaccact	tgacagcctt	gacatgtgtc	atctggattt	taacagacag	34860
atatcattga	gatattctgac	tcacacctga	aaacactgta	actagaacta	ctctctctct	34920
ctctctctct	ttctctctct	ctctctctct	ctctctctct	ctctctttat	gcctaaaaca	34980
gagctatgtg	gaggctaagc	aggacttcca	aaatttgaaa	caaataagga	acctgttcaa	35040
agaaggcaga	ttgaggcac	ttgaaagaaa	ttgcctggca	gggtgtcatg	actccattag	35100
ggggaacgta	ccatttgaag	ccaagagaat	aaagaagtca	tggatgacta	gcttgagaag	35160
gcgtaactgg	agtcagggaa	acagttactg	agagagatgt	cctgcaatgt	agggaaaaag	35220
agggagcttc	cccaaactca	acacaatgct	ctatgaaaca	cacatcagtt	gggatgttca	35280
ctaaactcag	aggatcacaa	tgacagatta	taagagcgcc	agtcaagtta	gaacttctta	35340
ttccatctta	cgaccttcc	cttctctccc	tacactcctc	tgtattcttc	ctcttgctcc	35400
cccacagcct	ctatactctg	agaagcagtc	tcttgctctg	ctccttctgt	gaattgttaa	35460
ttcaattatt	tcaaatttct	cacttcattc	tctgtctagt	gggtgtgagt	cctgggtgtc	35520
tggggatctc	atatgtgatt	tattgtggtc	agaatgcaaa	aatataggca	aaatcataac	35580
atagtagtta	aattttataat	tatttggatt	attgaaatac	atctataaaa	agtattttat	35640
aatttgtttt	gaatactttg	tctttggaga	ggccagaaat	cacagacagg	aaaggaggag	35700
gagatagaga	gaggagaaga	tggggaaacc	aacctggcct	ctcctcctac	tcctggattt	35760
ccaatctgat	aacagagttg	tgagaaaaga	cagagcttaa	gtttacatgg	ataagacttt	35820
tgactagact	ttttactacc	ttggaaataa	gcttttttaa	aaaatttatt	acctgttggt	35880
aatcagaaaa	gccaagagag	aactactcag	agttttatta	gttagaggaa	cctggggatt	35940
gagggtgagt	actcaaaatg	tcaggtttgg	aggggcagaa	gcaaaattta	aaaaattgtt	36000
tatacctgca	acctgtcaag	tccagtttac	acaataatac	aaaataaacc	aatttaactt	36060
gaacacataa	aatgttttca	tgatgtggat	tcaatttaac	aagggcacaa	catgatacca	36120
cagtcatctt	tcaaaattct	catcttttag	attcagcttt	ggtaccttaa	agatatattg	36180
ggttacagca	acatttctgc	aggaagctaa	tatgtattcc	ttaaaaagag	ttccatgatc	36240
aaagatatattg	gtgcacaaact	tgttcactca	aaaaataaat	agaggccagc	catggtggct	36300
catgcctgta	atcccagaac	tttggcagga	caaggcagga	ggatcacttg	agccccagag	36360
ttggaaacca	gcctagacaa	tatagcaaga	ccctgtctct	aaaaaagaaa	tagctgaaag	36420
tgggtggcatg	tgctgtagt	cccagctact	tgggaggctg	aggtgtgatg	atcatctgag	36480
tccaggaatt	taaagctgct	gtgaactatg	aaggctccat	tgaccccccag	cctgggtgac	36540
agagcaagac	cctacctcaa	aaagtaaaat	aaataaataa	gtaagtaggg	tttttttttg	36600
ttgttttttg	tttttactg	taggctcct	ctgggtgtgg	agaatgtttt	aatgtgtgtt	36660
gtaagggaac	aattctataa	gaaagaatac	ctgggtgtgg	cattttcaag	ctcattttgt	36720
caaaggctgt	tattattccc	aaatatttag	gcttccattt	attaaatatt	tgctgaacat	36780
ctaccatgtg	caggcatggg	ggatgggtgt	tggggtttta	ttaaatata	caaatgggct	36840
cctctcactc	aaagatctta	tgatctaata	gattgcttat	tttagtcctc	atcttagttt	36900
gaaaatctct	ttggagagat	ctgttctatg	ccatttttca	gatttctgcc	actgcaaaac	36960

tatttgttaa	cacaaaaata	caaattacct	gaggaagaag	ttgtgttgta	aaatagatgt	37020
gtttttataa	aacaaaattt	aaaagtgcct	tcacaacca	gcataacaaa	tatgcttttg	37080
tcacctttcc	ttctgtcctt	tgatcaactt	cagctgatta	tggaaaattc	ttgataaagt	37140
tataaagtca	tccacctctt	ttttctctag	gtctaactcc	tttccttgta	ttggacgtca	37200
aatccccaca	ttctactctt	tctgtccctt	cagaaaagat	ctgacatcag	tcattcggat	37260
tggtgtgggc	ttcctgtgag	atattttgac	aaagggtttc	tttatgtgtt	agcactgaac	37320
agttttggca	attttcttcc	ccttcaccag	cctactacac	atgaccacac	aggtgtttac	37380
atatgtacat	acacacacgt	gcacacttac	cttccttcaa	aacaggccaa	atgcaacatc	37440
ttgtctcctt	cttgaggca	aagtaggaga	cctgtgttga	gtgacagaat	tcctttacct	37500
tctgggctat	agcatcccag	agaatgtcct	cactgggag	ttttctatgc	atggctttcc	37560
ttgattccca	ggaagatcta	catccagtta	acttaactat	ggcagctgaa	tcctgacaat	37620
ggtgtgaagg	aaaacatggt	ggatacaaat	gagcaccatg	ggtacagctt	gatgatgtcc	37680
agctttatgc	ataatgtgga	ctggcagggc	ctttgatcat	gcctcaatag	ctgtgaagat	37740
agcagatggc	aaaagaagtg	gtaggttcta	agtaggttcc	ccttctctac	tggttgccac	37800
agactcaaaa	ggaattttta	ctacaggaag	ggacccaaat	attctagcct	agcctctata	37860
ctctgaaaaa	cagtctcttg	ctctactccc	tctgtgaatt	gttaattcaa	ttctttcaaa	37920
tttctcaatg	gattctctgt	ctagtgggtg	tgggtcctgg	gtgtctgggg	atctcatata	37980
tgatttcttg	cggtaaaaat	acaaaaatgt	aggcaaaatt	ataacatagt	agttgaattt	38040
ataattattd	ggattattga	agtacatcca	taaaaagtat	tttataagtt	gttttaaggc	38100
cttgacctgt	tcaacatatt	tcttaaatgt	atctcaattc	cttctacatt	ttgctatcat	38160
acgtatcgcc	atttatgtct	aggctaccat	catctaagac	cacaataata	ggtaccaact	38220
tatcttccca	cttttgtcct	ctgcaattga	aaaataaatc	agggttttta	cttctgacca	38280
aatggagggt	ctcagggtgtt	tgactaaagg	tagtatagca	ctgtgatgct	aaaataaggg	38340
aaagcaatga	cgtgacattg	agtgattgtg	agtggccaac	catcccatt	ctcctagtgt	38400
gcttgggact	gagggagttc	ccagaacatg	agactttcag	tgctaccact	gaaaatgtcc	38460
tgcccaaac	aggataagtt	gttcatctga	aagccctgtg	tttttaccac	ttgacagctt	38520
gtagaaagtt	tccagtctgt	agtacaggaa	ggagagactc	ttatatagct	tgccaggctt	38580
gagttgagga	gatagaagca	aaaattttaga	gagacaagga	ggctagaatt	tgaagggtga	38640
gtagtaaaag	gaagaaagct	gcacaaaagag	ggagacagta	gactacagat	gtgcagaagg	38700
gtccctttca	gttaatttag	ttctgatctc	accattcatg	agagatgtct	tagtctgttc	38760
tgtgtttcta	ttaatggaata	gctgacactg	ggttaattat	caagaaaaga	ggtttatttc	38820
aatcatgctt	ctggagggtt	ggaagtccac	aaagcatagc	accagaatcc	acttggcttc	38880
tcgtgagggc	ctgtgctgca	atagaacatg	acaagaagta	gaaggggaag	cagacagaga	38940
tcaaacacaa	gaggcaccct	actctcctaa	gaactaatcc	atccctgtga	gagctaattc	39000
aatcttgtga	gagcaagatg	gcacttctca	caatcttgtg	agagtaagaa	cccactgacc	39060
accgtgagaa	ctgcaccaag	cttcccattg	aggcagaacc	ctcatgaccc	aaatgtctcc	39120
cattagattc	cacatcttaa	agttccatct	cctggcattg	ctgcaactgg	aattaagggt	39180
ccaacatgag	ttgtggagga	gacactcaaa	cctgtagcaag	aggcaactat	ctgagactgg	39240
ggaaagaaat	atcagaaaaa	agtaagtga	aaacccatgg	accctcacac	agggctgcaa	39300
ttagtttgtg	atcccaaaag	aaggaaagga	aattgtcata	atccaataaa	cattgggtag	39360
aatattcaga	ggagccctgt	gttagtaatg	aggataaatt	agccttaggc	taaagactat	39420
tctggaccct	tcctaacaaa	gcttacagga	taggatccaa	ttgcttccaa	ataactgtgt	39480
ttcagaaaaa	caaaaaacaa	acaaaaaaa	actaacatta	cttaagaaa	tataacaaaa	39540
tccataaatc	aggagcaag	tcaagaaata	gaaaaagaat	cagaaatgat	ggagatgtcg	39600
gaactaacag	acaaaaatct	aaaactatta	taaatattcc	ctacatgtcc	agtaaggtag	39660
aggaaaagat	gagcataata	aagagaaaaa	ttgaagatat	taaaaaatac	tcaaatggaa	39720
cttctagaga	aaacaattat	aatgagtga	tcaaaaaata	aactacaggg	ttttatagaa	39780
aattggacac	tgtaaaagaa	atattggtga	acttgaacat	aaagcaaaag	aaaatgaccc	39840
caaacaaaca	ctgagagtaa	aaaaaaaaa	gacaaaaaac	ccccagaaaa	tcagtggagt	39900
atgggaaaaat	ataattgtta	taacctacaa	gtaattgaag	ttcaggagga	gaagtggaga	39960
atggaatatt	ttgaaataat	gatgactaaa	agttccaaaa	tttgatgaaa	atgataaccc	40020
aaagactcaa	gaatttccat	gaacctcaag	cagaagaaag	gcttttaagc	aataaaggaa	40080
gctaagaaaa	atcatagtaa	aattggcagg	agagtgggaa	gaacagtga	aaaaacattt	40140
tttaagagca	accagagaaa	aaagataggt	tacatggcaa	ggaacaaagg	tgagaataat	40200
gccatccttt	ttgtcagaaa	ctattttaagc	caaataataa	taggcatctt	taaaaataaa	40260
aaatgaaaaa	aaatttgaat	aagaattgtt	tatacaatga	aaatatcttt	caaaaatcaa	40320
gtcaaaagaca	gattttcaga	caacaaaagg	aggagatgac	tcataatcac	caagtgtgca	40380
cttaaacaaa	tattaaaata	aattatctag	gcagatgaaa	aatgataaca	ggtggaaaaa	40440
ttgggcctac	acaaaggaat	gaagagagaa	atgaataata	tgtagataaa	tgtagaacag	40500
attttttatt	cttttccagc	ctctttaaga	gatgaataac	tctttaaagt	taaaaacaaa	40560
aaagaacaaa	ttctggagtt	tgtgtcatat	atagaagtaa	aaagcatggc	aacctagcac	40620


```

aaaagggttaa ataggaaatg gatggatact gatgcaaggt ccatatacca tgagtgaagt 40680
cgtataatat tacatgaagg tgaactgtga taaattttaa atgcatatgt actgtaaacc 40740
tcagaagaac cactaaaaat agttatgtct aacaagacaa tagtaaaaaat aaaataaagt 40800
aataaaaaata actaataaaaa ataataaacag ggagagagga ataaaaaaca aagagaagac 40860
aagtaaatag aaaacaaata accaatatag aagattttaa accaacagta tcaatagcta 40920
cactgcatgt taacagtcta gacacttgaa ttaaattggca gatactggcg ggattggatt 40980
aaaaggcaat acccaactat atactgcctg aaatatagta cttaccaagt ggatttttaa 41040
aattaataat aaaagaatgg caaaagatat accatgcaaa cactaaccaa aagaaagctg 41100
gagtaactat attaataatat gacatcttat agttccaata tttatagttg gaaaattaac 41160
acctcaatct cagtaattga tagcaaaagt agacagaaaa atcagtgaag atacagaaga 41220
cttgaataac actgtcaatc aacttgacct aattaacatt tatagaacac tccatataac 41280
aagaggaaaa tctacattct ttccaaggctc attcagtaag atagacaata tcctgggtca 41340
taaaacaagg ctcaatcaat ttaagagaat taatgtaata caaagtatat tctctgacca 41400
caggagaatt aaactagaac tcaataacgc aaatatatct gcagaaaact ccaaataattt 41460
ggaaacccaa ccataggtca acgaggaaat cacaatgat atcagaaaag attttccctg 41520
aatgaaaaca aaacacaata tatcaaaaat tgtggaatgc aactaaagca gtgcttagag 41580
aatttataatc aattaaattt tagttagaa agaaggctt tcttaaatca atgatctcag 41640
cttccacctt agaaaactag aaaaagaaga ggaatttaag caaaagttaa aaataaaaaat 41700
aaaagctacc tttattgaac acttaccagg cattatctta agtgatttat gctaagtgtc 41760
ttttccaatc tggtcataac attatgagtt agtcatgatt aatatccac atttcagatg 41820
aggaaattaa atcttgctga acttctgtaa cttccaagtt tatgtagtta accaagagcc 41880
agagttatta tttgaactca ggtctctatg ctggttgtaac ttcaacatcc tacagattga 41940
gtgtatttat accgttaagc atattttatt atcttcagggt ttgatatttc tttgttgcat 42000
atatattata catatatata aatactatat acacatacat acacgaattc ttcatttacc 42060
tagaaccaaa aggaccttct cctagcatgc tggagaatgt ctagagacca gacattgcta 42120
ttatcatctc tgaaaactgt ggcatcagtg aaatgacctc ccatgaagga tgtcactggg 42180
aagtctgtga tggcagtgca tgaataattt gcacttgctc ctgccagctc ttcaagctac 42240
atgatctcga gtcattggaga tatgggtgag caactgagta tcacagtgtc ctccatggcc 42300
accataactc actcaagtct taatatcttc tcgcaaggat aaaccagacc tcctttaaacc 42360
aaaaataagc tcatgttctt tacatgagaa gccaaattca tctttctcca agcacaaggt 42420
agtaaagagt gctggacatc tgatttctat ttgtagaact gccaggcaat tattgtgtgt 42480
cttgggcaag ggaacttgatt ttctaaatcc tcagttttct ctggttgcaa atggaagcaa 42540
aaagatctct ccaatgtcct ctcttggtt gctgctttct gcacttttct tcctacacaa 42600
ttgacaattt taactgcctt atcggtataa ctactgaaga atcagttccc aatatccatg 42660
aaggtaaaag tcagatatta ttattcatcc actccaaatc ctccaatgac ttcccatttc 42720
actaaaataa aaatgtaagt ccttacgatg tcttcacaac atagtttctc ccatccctct 42780
gtacagccta ttttttctt ggccttatct attattcctt cctctgtat tagtcagctg 42840
ggttcccggt atgaaatacc actgactgag gagcttcaac aacacaaatg tattttatta 42900
ttttttttat tttttttat tttttttat ttttgagaca aagtctcacc ctggtgcccc 42960
aactggagtg cagtggcaaa atcttggtc actacaacct ccacctcca ggttcaagca 43020
attctcctac ctccagctcc cgagtagctg ggactacagg catgcaccac catgcccggc 43080
taatttttgt atttttagta gagatggagt ttctgtatgt tggccaggct ggtcttgaa 43140
tctgatctg atgatctgcc cacctcagcc tcccaaagt ctgggattac aggcgtgagc 43200
caccatgcct ggccacaaat ttattttctc acagtcctag aggcgtgaag tctaaaatca 43260
aggtgtcagc agatttctatt tcttctgacg cctctctcct tgactttag atgactgcct 43320
tcttgatgtg tctttacatg gtctttctc tgtgcacgca catacctggt atctctgtgc 43380
atgcccacaa tcccttttct tataaggaca tcagtcagat tggatgaggg cccaacctga 43440
aggcctcatt ttaacttaat cacatcgtaa aaggccctat ctccaaggac tcttagatat 43500
attagaggtt agggcttcaa tgtgtgaatt tgataagggg agggacataa ttcagcccat 43560
aatactctca ctactgtgt tccagctaaa ctggtcttct gcattggtctt caaacacact 43620
cttgtctcaa ggcttttgca ctggctattc cttctccctg tgccgctgca aagaacgata 43680
atgtttgccc caatcttgta gttttaagga ccttttcata acctaagca catgtcaacc 43740
catggccaga accctatgta agatttatcc ccaattttt ttttgccctc cttgtctctc 43800
tcttttaatt tgttcccttt ggcttcctta agtattaact ctaacagata acagctttac 43860
agtttggtt ccctgtttt attcaactca gtttcagggt attataatgt cttcatttcc 43920
gatgtggccg ccatattgga tgggtgctgaa tcacccattc aaaagaaatg tgttaatgcc 43980
agtttatgtg ccagtcctcg ttagtatctg gaaatacaca ggaaaatatt agggtagga 44040
ttaatgaaga ctaatgaatt cagtagtggt atacacagag tatcatcaga gcccaaaaa 44100
ggaacattta tatgaaagaa tatcaggaag gcgtctggaa agagtggta ctacagctac 44160
agtttgaaaa atgagcaact tagccaggtat tacagagggg aagagaattt tctagactct 44220
gttagcatca tgcattgaaag tatggaagca tgaaacagta tgatacattt acagaaaagc 44280

```

```

gcatgtagag aagtatgaga taatggggct tgattgatag aaggctagag aataaaaatc 44340
atgtgtgcta agataatttta tcttgaaagg tccaggggagc attttaaaat tcttaacaag 44400
ggaatcaattt ggtcggtatct ccaatttaga acggttagta ttccatcatt gaagacagtg 44460
acatgtaaga gtgagttctg gatgaaggaa gccaggtagg gagttactgt agtaattctt 44520
ctatgaagaa atgaggacta gaacaagagc agaggcaaca caaatataat ggaaggtagc 44580
gctatatgac acagtaagga aattgaatct atagaattgg tgctgggttt aaacatgaga 44640
attgaacgag aagaaaactc taaaataaat cccagggttc tgggtttggc aaagaaagag 44700
atgatgatat tattcatcat tagaaggagg agaaacatta acttgacagc acattgagag 44760
tagtgaaact tccgtaagag ctgtccagtg aacagggtatt ttatacatt ttgtcaaggg 44820
caagaaaatt tcccattgat tggggcctat ttagtgcct tcttcttctt gcaagagtag 44880
tatttttctc aacacacctt gtaatcttag gcaagggaag gaaggctttt atttggagtt 44940
gcttttataat gagaccaagt ataactaaaa agatatagca tctctcaac cccaccagc 45000
ccccattcat tgcagcacct cctaagggtg agatgtcccc tgtatcccaa aaacgtgagt 45060
ggtgactaag tgtgatgtgg cacaacggt gaccacacta aagtgaccac atagtacttc 45120
gctgctccca ctgggaaaca aaacataata caaatataca atgcttcaca gagtagcagg 45180
tgattctttt ttttttctt aaagtcagat ttattgaggt atagtttaca ctaaacagtt 45240
tactttttgt aactttttaag ttcagaggca tcttttgacc tcaaaaacct agttaagtaa 45300
gcaaaggaga actttttttt tgttttctt tccatgatga ggaagctgat gattaaagag 45360
gttaaattgac ttgtcccggt tcatacttag tcttgagtaa gcaagattaa gtcttctgac 45420
tcttaattct gctttttcca agatatacaca gaggatcttc cagaacagct tacttagaaa 45480
ggacataata tctgttcctc ataacagttt tatggtaata ggaatttaat cttctaagta 45540
atcaactcag gcttttccac ttaccagtg aattgcagga ggaaactata gaacttgaaa 45600
cactatcttc cacatttctg tattgatcct gtttgatggt ttgtgatatt aagggaatac 45660
atgactatca tctagtgttt caagaccaag ccttcaaag tagagtaagc cacaacataa 45720
aaagtgttta gttaccttct tggtctaga gagtcattaa atccaacctt tctcttctgt 45780
ctagctgcct aggagtggct tctgttttaa tcatcatcat cattatttgt attagtctgt 45840
tctcgattg ctataaagaa ctacccgaga ctgggtaatt taaaaagaa agagatttaa 45900
ttgactcaca gttccacagg ctgtacaggg agcatggata tggaggcatc aggaaacaca 45960
atcatgggtg aagggtgaagg ggaagtaggc acatcttcac atggcagagc aagagagaca 46020
gagtgaagg gaaggttaaca cttttaaaca accagatctc agccgggcgt ggtggctcac 46080
acttgcaatc cttagcattt gggaggccaa ggtgagtaga ttgcctgagc tcaggagtct 46140
gagaccagcc tgggcagcat ggtaaaaact cgtgtctact aaaatacaaa aaattagccg 46200
agagcggtg cacatacctg ttatcccagc tactcggtat gctgaggcac aagaattgct 46260
tgaacccagg aggcagaggt tgcagtgagt tgacattgtg ctactgcact ccagcctggg 46320
taacagggtg agactctctg tttccaaaat atatatacac acacacacac atacacacac 46380
acacacacac acacacatac atatatttat ttatttattt atttatttat ttatttgtgt 46440
gtgtgtgtgt gtgtgtatgt gtgtctgtat gtgtatgtgt atatatataa 46500
acaaccagat ctctgtagaa ctctataatg agaccgcacc agggagatgg tgctaaacta 46560
ttagaaacca ccccatgat tgaatcacct cccgccaggt cccacctcca acaccgcagg 46620
ttacaattca acatgagatt tgagtgggga cacagagcca aaccatacca ctattttatc 46680
atttctatta agagttccct gcatttcaat ctggcctaga atacctgctg agatcataat 46740
cttatgggta ttaagtgata gtgtcataaa ttaccatctg caaaatacta tttcagaagc 46800
tgtaattctc tgtcaagttc tctgttatat tgctacactc tgcagatttt tagaaagtta 46860
aatattatat atgtataact ctgcattaat gcaatttatt tgattcttgt ggtatctcag 46920
gaatgtaata tgcaaaagtt gtatagatgt ttttctttag tgtaatgaat tattaggata 46980
tgactctttt gtaaaaggga gaaagaacag taatttgaat ccctgcagag atagcaatgt 47040
atttgagagc aatttgaaat gttttatctt tatctactaa aaattccagg ttgaaaagaa 47100
taaaatcacc tctctctctc aaatgcttta ctctgtagt attattatgt aattattcac 47160
aaaaagggtg gttcttggtt cacaatcact accttctcct tggagatttc aaaaaataat 47220
agatgagaaa ttatgtccct caaaaaagac cacaaaagtt tatccagaaa actttccctg 47280
tattgatttt tctttaataa atgaagctac atgttttcat ttttaagaaa tataaaacct 47340
taacttggtc ttctctcttg atacagttta ggaggtcttg catgcattct ctaacctcac 47400
ctccttcccc tttctccaac tacactgtcc tctttcagtt cactgaacac tccaagctca 47460
ttcatgcctc agggcctttg catatccaga tatctccttc tgctcactcc ttattaccac 47520
catggtccat ttgaactgga tcctcccagt tttttctcta tcacttcacc ctattcattt 47580
ccttattaca atttacaaaa accaaaaagt gttactctt atctctcttg ccagctctat 47640
cagagctgct accatactg tgttgatgt catgtataac cagcggctag ttccatgcct 47700
ggtacatcat tgcttaatat cttttgttg aatgaatgaa tgaatggatg aatgacaaaa 47760
attcctaatac aaatttctg ttgctacctt ttatacttac aacttaatat tacatcctta 47820
tactctaaac tcatacatat ttctataaat aggagaaact ttacctacag aaatgcaca 47880
ctaagattgt aagcaggaga aagataactg aagttttctg aatttagaat ctattctggt 47940

```

cacttttctaa	tgagttccag	aatctatggt	cccatatctt	tggaagcaa	ataattgact	48000
gttaatatgct	tttcaaatat	gtcatcagaa	tggcagatga	ggcctgaaat	tctcaatgga	48060
ttgacaaaaa	gcccctgtcc	tcttggtaga	atgaaatata	gtacttgtaa	cacttcctta	48120
cttcccccaa	attactttatt	ctttaagtta	cagaaaagga	attccacaaa	ctacaatgaa	48180
aactcaaagg	aagtaattac	tttaggtaat	atacaagtta	atactatgtc	ttttgtttgg	48240
tttttcttca	gacagaaaat	aacccaaact	ttttttttta	ttcattcttg	aattttagaa	48300
accattgtta	aagtattaat	catttctgag	ccatccgtag	attgagttaa	ttcatccttt	48360
ctttaagttc	ctctcaacaa	acttagccat	gttttatgat	gtataaaaca	tgctaaaata	48420
tttttaatat	aaacccaaat	ggatttttgg	caggtagtgc	gaccctgagc	agtcagtata	48480
tgacagggag	accaaaacaa	cccaaccagt	tctgatttta	cacaagttcc	ttgccttctc	48540
ctttaacata	tatcggaatt	gtacacctag	atgggttcagt	gttatgggaa	gttgacagac	48600
agaaaatgca	gtattatatt	aggtcaaagg	ttgacaaaact	tttctgttaa	agtgaatatt	48660
tcagactttg	tgagacatgt	actcgctgtc	acaatgacac	aactaagcca	tcttagaggg	48720
aaagcaatca	cagacgatat	ataggtgaat	atgcataagc	aggtgcctta	cttcaggctg	48780
ctgtcacaaa	gtaccataga	ctgggtggct	tacaaacata	tttctcactg	ttctggaagc	48840
tgaggagtcc	taggtcagag	tgtctctctt	ctcagttaca	gactgctgtc	ttcttgtatc	48900
ctcacatggc	aagaagagga	caagagagct	ctctggagtc	cctttcatga	ggacattaat	48960
cccattcatg	agggctaaaa	cctcataccc	taatcacctc	ccaaaggctc	taccacctaa	49020
taagatcaca	ctgggtgcta	aaatttcaac	atatgaattt	tggaagaac	acaaacattc	49080
agtcactgct	agtgtgctat	aataaaactt	gatttacaaa	agtacactgt	aggccagaca	49140
tggttgctcag	accatagctc	gacaaccttt	gaattaggtg	atcactacaa	catgtaaatg	49200
aatatgcatt	gtgtcttagt	tcaggctgct	ataacaaagt	accatagaca	tggtgggttta	49260
caaatatggt	tctcacagtt	ctggaaacta	gggagtccta	agtcagagag	tccctcttcc	49320
cagatataga	ctattgtctt	tttcttgtat	cctcacatgg	caaaaagagg	ataagagcac	49380
gatacctaga	tgtattttgt	atggggagtc	agccatttta	acattttatta	aacacccact	49440
aaactccaag	gtttccagac	agtgtcataa	ataagcttca	caactgatct	gtattcagta	49500
gatgtgaaaa	ttgaaactgc	ctgggttatta	atctgtgtgg	ccaggccatg	aaccattctt	49560
cttcggtcag	gttcagccct	taagcacaca	tgccacagag	tcaaaaacaa	cctaggtttg	49620
aatcttgctc	agctattttac	tgatgggtgt	ctgtatggaa	aattactaga	cctctgtgta	49680
tctcagaatt	ttccoctatc	tcatggagtg	gctaggagga	ttaagaaata	aaccacaaat	49740
agcactcgct	tagaacagtg	ctctctgtct	aacgcagtgt	gctagctttg	accacaattc	49800
tgacagacact	tcagggtcca	gagagttctt	tcagaagttc	cacaaatggt	tgatttaaat	49860
gctgctaaaa	aatatttttc	actacttaaa	ataccttaat	aaaaagaaac	atatattcat	49920
gcaatgtata	ccaaattgaa	cacatattct	cattctatat	gttaacttgt	gttgtcagtt	49980
aacttttaca	cagacattga	agcaaagctt	ttcctgttct	tcctgtttat	ttaaaaattt	50040
tgacagagaat	tcaaggtaag	ctgataaaaa	taattcttac	cattgtaacc	acttagctgt	50100
gtgaatcaga	attgtcttga	cactgtgcag	ctgaagtaaa	aacaatctaa	aattggaggc	50160
tgagggtggac	ataagattct	catctctaatt	attgtgttaa	tcagaaaaac	tttattattc	50220
ccattgattg	gctttacaat	aagtaaatgt	tataaatcgg	ttttaaagcc	tttatatgaa	50280
aatgactcac	tttttctggt	ttatatattt	aaggtttcat	gtaaattttc	atttgagact	50340
aaaaagaca	tttttcagca	gaaacattta	actagtactt	aactagtcaa	gatctgcctc	50400
tttagattgg	ggcttttttt	agatcaagac	ttgattctta	cgtatctttt	ttaaaaatta	50460
gactctgata	attaggtcaa	gacaaagagt	tacccaagca	gagttcttca	ttctccttta	50520
ctctcctacc	cacatcatta	tgtcctagtg	gctttaatag	atattttctc	cttagatctt	50580
tataataatc	ctctgagttt	gccaggcagg	tgatattagt	ctcataaatg	aagaagggtg	50640
atgagactca	aagaatataa	atgactccca	agggcatatg	actaaacaag	tggtagggtc	50700
aagactcaaa	tcagaaggct	tctgactact	aatcgagagt	ttttcccacc	cgttttgatg	50760
tcaaagatgg	ctgagatttt	ggggaggggc	tttgggagtt	cagagggtgac	ccttttagcca	50820
ctgctgaagc	tgatagagat	gcaagatcag	taaacatatc	ttattctttg	tttgtttttc	50880
catagggttg	ctaggctttc	tagtgattaa	aaaaagcaat	gtagacactt	gtaaagtgtt	50940
ttaatattat	gtcctggact	ctgtctcact	tgttgcaaga	gtcatttgac	atgacgcctt	51000
gctaagaaat	ccacacgtag	gtctcaaatt	cagtcctggg	tgatttctta	gggtctatgc	51060
tccatgaggg	cagggacagt	actagatatt	tgtgtccaca	aacatctaga	ggcatccaat	51120
tcttatctgt	tcataaatat	atgatagaat	caaccatcat	tctatacatt	tttcttttag	51180
tcttttttcc	ttttgtctacc	attgcttttt	ggaccattat	tgcaggcata	tcataaaactt	51240
ccttggtaac	tcttggtcatt	ttggtcctgc	ctccctcaga	tttttttctg	cataatctca	51300
aagcatccag	caatgcttct	ccagcgcctc	tcagaacaag	gccagagta	tctcctggg	51360
attcaagttt	cttcatgata	tgggcatgcc	ttatgaatac	accttcgctc	tggtctcatc	51420
accgtccctg	cagtcctctc	tctatatggg	acatgctcat	tgtgtctttg	tgtctctgcc	51480
gttgcttgct	tggttgccag	ctttcatttc	tctacacaga	tccaaattgt	acccatgatt	51540
tctaccta	ttcaactcta	attttctatg	atccctgccc	cagtaatttt	agcccataag	51600

aatTTTTata	gcacaaattt	tgtacttaat	gacacttgta	ttatttcatg	ggtagttag	51660
tgcaatctct	atgaggctgt	aggcaatcaa	gtaagatatt	taaattgtac	ccacattgtc	51720
ctaccttccc	acactgtaga	cacctactaa	attggagtaa	tgtgtaataa	aataatttat	51780
agattttaagt	aattgaaatc	atcccaaggc	acagagagaa	caaacaaaca	tattccttag	51840
acatgtctaa	ttcatgtgcc	atggttatct	tgagataaatt	ttccactgta	cttttaataa	51900
aagagacacc	ttaaattat	catagtttgc	ttttgttttt	cctccttgag	cattgccatg	51960
gttattgata	tttattaaat	caaactctgcc	actgaagcat	atataaccac	tcaagaatga	52020
attacagcaa	atatttaact	gcctgcctgc	ctttgaattg	gttaatcaga	atgagtgatg	52080
cattagctgg	ttcaactcct	ttcacctgga	aagatggcta	togattccct	ttattgtgat	52140
ttaaaagatt	tgaggaaagt	ctaaaagcat	ttcaaaaagt	ttcaataagt	aagacataat	52200
tctgcttttc	ataacacta	aatgaaaaaa	aaaactaaat	acaaaaaaga	ttcttcgagt	52260
aaaattctgt	cattctccac	tttttattgt	gttaaaaaaa	tcaactataa	tatgagtga	52320
agtcaagttc	tgctggtatg	tatgaaatga	gaagagaggt	gctgtaggtt	attaataaca	52380
gggatgtttt	tgtacctgtt	ctggaagatg	aatatgcatt	tttgtgttgg	agccagtcca	52440
taagtcaaat	tcttttaaaa	caggcagagg	ccagccttgt	tctggctgca	gcctgcacgg	52500
ccagaaaaac	agaggaagaa	gcgccttcca	ccagcatgaa	atggtatcca	ctttgatatg	52560
taaacattct	tccttcacct	gaagttgatg	tcagggtaac	taaaacacaa	gaacttggtt	52620
gactgagcag	aagcaatctg	ttgtattctt	ctgtttttca	aaaccagaga	taatcaagga	52680
ctgttatgtc	agatgctagc	tgacctaaact	tttgagtgga	aaacgaacca	ggaatctgag	52740
ataactggat	agtctccgca	tatttttaaa	aaatcagcat	gtctcttgaa	actgctcaga	52800
cttagctgat	ttttgtagtt	gttgtgtgtc	tctttatagt	ttttttctt	ttttctctgt	52860
attccagact	cctggcattt	aatctattgg	gtgggaataa	atgacgtaaa	ttgtactcat	52920
atttgcaaaa	agcccaaaaa	agatgttaag	gctgttaaca	acagacgtct	acgctacagt	52980
tgaaatgggc	taggtgaatt	gagacagaat	tctatgaaca	ttttatgaag	ctttccttaa	53040
atttctaatt	gaaaattgct	actgtttgtg	ttactgttga	ttcatttcat	ttaagtcttg	53100
gtcttgtaga	ttctattgat	acccattctt	tggggatgaa	aaacaatgtt	aatgctgcac	53160
atttgttagg	aaaatgggct	gggtgggaga	cggagggtct	tggccgccat	cctagggcat	53220
tcctgtgcat	tgcatgtgtc	agcatttctg	tatctaaaca	tcgaaaaagt	acaataaaaa	53280
tacagtatta	taatctggga	ccactatcat	acacgcagtc	tttcttgacc	aaaacatcct	53340
tatgccgtgc	gtggctgtgc	tgagaaggag	gagcactcgt	cctctccaag	cacatccaat	53400
ttgcctcagt	cccgccctgt	taccagaact	gagctcattt	gttattgaag	53460	
ttgcttttct	tatacccgagc	ccagttttcc	attgatgctt	ccgcttattc	tatgtaggtg	53520
ttttcattta	tcacacaata	tccttagtct	aagacttcat	ttgaggaaag	aaaaaaaaaa	53580
aaacctctgc	tgctaaaaat	aatttgaaaa	ctcctagaca	ataggacctt	tcaaattccct	53640
tacaattcta	acaaccagtg	aattcatgct	tcacagccat	tcagatctat	acgtatgaaa	53700
tacaagacat	atgtgttcaa	ctttccttca	tcgcctcctt	taaactttct	tttttctttt	53760
ttgagtcagt	cctgcaggga	acagaccaag	tgagcccat	aagagaacag	ttcattccca	53820
cctgtacact	cgagccagcc	aacaatctca	gcctcaccac	tccattcttg	attatgtaga	53880
accattccaa	tcctgagtga	caggatttgc	aaagacagct	gttgggctcc	acaaacattt	53940
ggaggagggg	aaaatgtgag	gttggttgc	acattctcct	acttcttttt	actgaaagga	54000
acagctgcga	tcttcacatg	taagatgaag	taaacaaaag	ctgacaatgc	ccagcaccat	54060
tcagcagtaa	gcattcaa	ggatttttcc	tcaccgcgtt	tgcaggcccg	gcacagtctc	54120
agttccccac	agcacagaga	ggaagcaagg	tctttgcatg	ccctgttcac	ttggccaac	54180
attgcccctag	atgtgcaatt	caaccatata	aaagcaagcg	taattcaaga	gccctgactc	54240
tttcaccatt	tcattgatggc	ttattactcg	cacacggcat	tgtcagggtt	cagcaactct	54300
gtgtgcagca	ggcagaaatg	cagtgtttat	ctctgacagg	caaacattct	taggctcatc	54360
ggcatgagta	tccacgctga	actggctatc	ggggaagtag	atccctgcat	gggaaagtca	54420
aggatgtttt	agagggactc	gatgattcaa	ccccattggc	tggttcattc	cttttgtttc	54480
ctctgccttt	ttatttgctt	acacagaaaa	cagagcatte	atttattttc	tttccctctc	54540
tcactctctat	taaccactag	ggccatctgc	ccactgaaa	ctgtctttgg	ctgcatgggt	54600
gtgaatccag	gtgacagtga	aaatgaagaa	gtgggtgattg	tccctgcatg	ttttctctga	54660
agggttaact	aaaaaaaaaa	aaaaaaaaag	cctttgtgtc	tgtttcgagt	tctccagttg	54720
aaatactgta	ctcattgaac	tgcaagaaat	aaaacctgca	agtaggcact	gtgctaattg	54780
aatagaaatg	tgtagacag	gggcactatt	tcttaaatat	gcttgaatta	agcaagttct	54840
tttaaaaagt	gggtgggact	tttgatgttt	tacatagtta	tattttcaat	ttcttgtgct	54900
cagtttaaaa	ctcaagattc	tttgtttctg	tgtgtgtcat	tagggcaggt	gttggtgatg	54960
aagtaaatgc	actgaaaaat	gatttgtgcc	ctaacctcca	tgaatttggg	55020	
gacttaatat	atccatacag	atctgacttt	taatcaaaag	tgtcacaggg	tggggtagg	55080
gataaggagg	aatttgtcta	caggatcctt	ttatcctaag	ttgccttaca	acatccaatt	55140
gtgttaccgt	tttttttttt	ttttttacaa	acctattgct	aggctaagaa	gtattgattc	55200
catctagtta	aacatgtcga	gtgagaggaa	gtgatcaatg	gattacttct	tcaaaacatt	55260

tatccccct	agtggaatcc	tctgattttt	ataacattaa	agagccaggt	ataggagttg	55320
ggagaagaat	ttaaaatttg	aaatgtcttc	ttttttcaac	tctgttcatg	aaaactctag	55380
ttaatgcctg	tgatttctcg	ggtctccctc	tctacttccc	tccattatca	ccaccttttg	55440
ttgctgctgg	ggctgtgggt	gccctgaagg	agcatgagat	ggccttcctc	agatatctgg	55500
gttttaaat	cagttcgtat	ctcttaagtg	aagtgggtag	ggccaaaaat	gactcactgc	55560
ccagtgggtgc	ctggaaacat	tatcagcaaa	gacattcacc	tagggagaaa	ataaacacag	55620
aattctcata	tgggcccatc	tttccaacca	tctcctggac	tttggtagtc	ttttctgagc	55680
cagatatctc	tctaaagtca	cttgctgtca	gacctttttg	tctcatattt	aggaatggct	55740
gaagttcttt	ataaatccca	gatgaaacag	gaggaaaact	actttgaaat	gatcatttcc	55800
atgatctcta	aaatgcaaca	aaccgtattt	cttaagatta	acggatgcga	gcctatgaca	55860
ggttcaatgt	gggaccaact	caagggaat	atgtgaccga	attattaaaa	ggcattcatt	55920
acatcgagct	ttcacctgaa	gaggggggtg	ccccagtag	tgcttctgca	gcagaattgc	55980
ttagaaactg	cagaacatcc	acaaggcaag	caccccaggc	cattgtgtct	gactgtcttt	56040
caaatagagg	aatggctagt	tgccatttgc	tgagggcatc	tcctgcaactg	agaactgtgt	56100
gcataatagg	atgatatcgc	cgagactcac	aagtatgcaa	gatagaaact	aggtgcattt	56160
ttttcagatt	aggaacttga	gaatggaaga	taaactgaga	aggggagcag	gggaggatag	56220
ggagagattg	ggcaacgggt	aagaagttac	agttaaatag	gaataaatcc	tggtgctcta	56280
ttgcacagca	gggtgcctgt	ggttgacaac	attgcattgt	ctatttcaaa	gtagatagaa	56340
gagaggcttt	tgaatgttct	caccacaagg	atatgataaa	tgtttgaggt	aatggatag	56400
ctaaatactc	tggtttgatt	attacatagg	gtatgccagt	ttcaaagcat	cacattttac	56460
cccataaata	tgtacaacta	tcatgtgtca	attaaaaata	aaattaaaaa	aaattttattg	56520
ataaaagaat	aaaaaatagc	atgctttttg	ccactctatt	tgataggcat	catttaaaaa	56580
acataagact	ataaggctgg	gcattgggtgc	acatgcacct	gtactcccag	ctatttagga	56640
ggccgagaca	agagggtgc	ttgagcccag	gagttcgaga	tcagcctgag	caacacagtg	56700
agacctcat	ctttaaaaaa	tatacgtagt	atagatgatg	taaaaggacc	acctttccaa	56760
atgagtcaat	ttcttcctta	atttaataca	ttcctaggtg	gtaagttttt	agagagctaa	56820
actaatcct	tgctatgtct	taggtttacca	tttataagtg	ttcttatggg	ggttagcttc	56880
attttatgta	ctaaacactt	atttgaaagg	atcgttgcct	ttttccatgt	gctaatttgt	56940
atagagagtt	tacgtattgt	gggaatttaa	tttattcctg	cattttatcc	aggaaaatag	57000
agccagatgc	ttttggacac	ttgttatttg	ctttaaataa	gatatcaggt	taagataaga	57060
aaaaatatatt	gattggatat	tgcgctcttc	tctggagaac	gtaaaacctt	tcacgtccta	57120
catctcatca	taccttacia	ctttagaaca	gactctaata	aaaggaaaat	gttttaactg	57180
tccttttact	tatagagggc	agggatggga	aagtgcactc	aattatacta	acagagctct	57240
caattaatct	aaattcaaag	cgaaggctcc	tggtgtggat	cagggaccct	gaataaacat	57300
gctttctagg	agacttcacc	gccagggtcc	aagagaggtc	acagtactcc	tgcaactcctg	57360
atggcagacc	cttggaattcc	attgcaggca	gtgacacacc	aagagtaaga	ggagctggct	57420
tcagccctgt	ccacgtgtgt	ctgctggctc	gcccttctgt	ctggcatctc	cacgcccctc	57480
ctcttcgga	ctcccttgca	tgaccttcca	aacagaatga	gctgttcttc	caaatatgac	57540
cctaactcat	tcaacagctg	gaggcaagag	caaactagag	ctcatgcaaa	gggtattagc	57600
attgagctct	ttttttcccc	ctctttctct	cagactaata	gatatttgga	aaattcatct	57660
aaaaaccttt	tttaaaaaaa	tttcattcat	aaaaatagcc	aaggggtttc	tcagagagtt	57720
tccttttagc	aagtcttttg	gaaggaaata	atgtctactt	tattattttt	ccttctctgtg	57780
cctgacctcg	gagaggatta	gaggggtctg	gcaaggagag	gagcctcaca	cccagccacc	57840
cacacttttc	cttctgacga	gacccaccgc	catgccatgc	tctgtacaag	ttctctggag	57900
gaacataaga	gccaccttat	gtgtgaggga	ataagaccat	gctttgaaga	gccagacttt	57960
ctccttttac	ttgctcttga	aattgttttc	actctgatct	gcaagtcaaa	gaaatgatct	58020
attatacatg	tagggacctc	aaaatctgtt	tgactctgtg	caaataaggt	acagtccctg	58080
ccttcttcct	atagttttct	tatcactagg	ggacttgaaa	aggatttagt	tcaacaactg	58140
agtcagcatg	agctcacaca	aaacaaatcg	tttcacacca	atgtcaattt	ctttctttct	58200
tcctcccccgc	ccccactttt	ttattttaat	aattcgggtc	gggaaatccc	acatgcttag	58260
tatcattttg	tcagcaaggt	acatgacaac	tctggtcaca	atagaaaagtc	gttttagtgaa	58320
ttaatgtcaa	ctctggaaaa	agtggctaata	gtgtcttggg	gagtttggtc	cttgccctgta	58380
cacctgcctg	aagatcacaa	gagtaagaag	gacaggtaac	acacactgaa	tgacaggatc	58440
agagggttagc	aatatctcag	aagcctgcc	tatgagccaa	acccaatata	aaacaattta	58500
gcaatattaa	gtaattatca	cattttgtac	ttcaataaaa	aatttaatat	catgactttg	58560
cggtaggaaa	ggccaagttt	aatgttgggt	aaaaatagtt	tggagattat	atttgtctgt	58620
cttcaatatg	atgtttcttt	caaaatatcc	gaattaaag	gaattaaag	tagaaataat	58680
atctaccatt	tattgagtac	ctactatctg	ctaaaataac	aaatttataa	tgaaggtgta	58740
actattccca	ttttacagat	gaaaaaaatg	gaggagggtta	agtaaccagc	actaagtcaa	58800
aatggctagc	aaagatcaaa	accaacatgt	tagccaatgt	taggcagttc	aagtcctttc	58860
ctctgtgcct	gaaatgaaag	aaattttcca	ggaattttgc	attggttatt	catatcaagt	58920

ataaaaaatg	ttattcattt	caatgaacca	aaattaagtc	aagacactta	aggactgtct	58980
agaggaaaaat	gaccaggata	ggaaagagtc	tggaagccat	gcaagaaaag	acacagatgg	59040
taaatatttta	acatgaaaaa	tagaagattt	gtgtgggata	taggagatac	cttcaaataa	59100
tgaagacctg	tggccaactg	gggcaaatat	tttaacccca	aataagttct	tcaacaataa	59160
cagtttcatt	gaaaataagg	gaagagaaaag	tgcccaaaat	gagcattctc	agtcactaga	59220
tgcagaatca	gaaagccacc	tgccagggtc	gttattgagt	gaaattctga	ctttgaaaac	59280
atttggccta	tacgagggat	tcttaactctg	atatccatgg	aagaacttta	ggtaaagtct	59340
ttgaatccct	ggaaattata	tgaaaaattg	tatatgtgtt	tttttttctt	gggaaatgg	59400
ggatatattt	cttcaaattcc	tcaaagcaat	tctacaatat	tggtggggca	gtgaaaggaa	59460
atcacttcca	attctacaat	tctgttttgt	aaaagtagtc	ttacccttca	cagaggtgca	59520
cccaagagaa	ttgtgtgatt	gatttaatat	taggggtgaa	tgaacataaa	caattgaaga	59580
aaacttcaag	gttttttagcc	ttgtaactaa	gatgttcata	caatttttaa	aaatctttta	59640
taaaagagaa	gcaaatttag	tggaagataa	ttaagtagaa	ataaataggc	aaactgaagc	59700
ctgtctaaga	ttccagtact	gagagtactg	agtacgtaat	tcaaataattc	ttttacatca	59760
tggcagaaaa	atgctgaaat	ctagggatat	atgctgacta	gtaactgagt	cttatataaa	59820
taattttttt	aaagccactt	gaactttaga	accataaaat	gttataaacac	taactgttct	59880
agagactact	tacttcagtc	tcctaattgtt	acaactgaag	aatcatgtc	tcagaaaaac	59940
taagagacat	attctggatt	ctggggcctg	gttccgtgatt	cacgtctctg	tctgcaaat	60000
taaggaaaaa	aaattctaaa	acattcaaat	ctggaataaa	cctatgcatt	tttcaagaag	60060
aaaatatata	tataaaaact	agtcttaact	acataataa	cacaaaggtc	cactttgtta	60120
tggatataagg	caagaccgta	aggaggtctt	atgcctcctt	atgaggcata	agattgcccc	60180
ttacgagttt	taaagtatga	attccacacg	tcagcatgct	ctgggatggt	ttcacagtat	60240
ttttaacatt	ggtcctcata	gtcttaattt	acctctggaa	atttccatgg	atgattaaac	60300
tatgaagaac	aaagaagcac	atggctgtgt	tatttagagt	ttagtgaatt	aggctctgca	60360
gtctgttggt	ttccttttag	atgttgctt	gtccgtttct	aagatcagag	ggagaccagc	60420
tccaggagtg	ctcatgaatg	catcagtggt	cttccatacc	atctgatgaa	aagattcatt	60480
ctgtgtctca	aaaggatatt	ccagatccag	ttctgagtca	gcaggctctt	ggtgtatata	60540
tatatatata	tatatatata	taaaattttt	cactgcttta	tttttttaaa	ttcatctgac	60600
aagggtcaagg	tgggattgag	acaagaaaaa	tttggaaacta	gtcccatattg	agacttattt	60660
ccttaaaaaa	ggaactgttt	tttaataatc	tgaaggttta	cataattttg	acatgcaagt	60720
tatatagctg	tgtgtgtacg	tttgcatgct	tgtgtgtatg	ttcatttttt	ttaaagagac	60780
tcaaactaaa	taaaattttg	attgggggtg	gtctgaaaag	gactatagtg	tcagctgtta	60840
taacttcagc	tttaaacact	aaacccaaac	atgggctaca	tgttccaatt	catactgaag	60900
ctttataata	tatggcctta	gagtagaagt	ataatgatct	cttgggagag	gcaagatcca	60960
attaaaattc	agtttgctgt	cttttggaag	gacataagca	gttatatatc	ctttgagacg	61020
gtcttaagaa	atacacacct	gagcttttaa	tggccactca	tttgactgtg	aagttgaatt	61080
tcaattcatc	tgggagtcag	agcacaagaa	agaattcaaa	taacaacact	ctgatatttc	61140
ctggtgtggt	accagtgaga	aaagagctcc	tggaaaatga	acgccttaca	catttgcaga	61200
tattaccaac	cacttataaa	aacaaggcta	ttgttatgca	taaaagtcac	tccttttagg	61260
ggaagcctaa	gagtgaattg	aatgtggct	gacatttcta	ccacaagata	aatgtttttt	61320
ttaatatcat	gtttaagttc	tcttagttaa	aaaagagaaa	agggaaaaaa	agaagaagaa	61380
agaaaaatat	ctataatcca	gtattcagaa	gcaattccag	acacttcacc	catgaaataa	61440
ctcactggaa	gacattacat	ttctaaccac	aaaagctatt	agcagccttt	tctaattctc	61500
ttttagttca	ataagggaat	tattgaccgt	attttgcatt	cccacatggt	tttaaaagac	61560
aaaaaacata	gtacattgaa	cagaaaacat	acgaatgctt	tctatttaat	ttattagttc	61620
atgactaaga	tgaaacactc	caacatatac	aataacctcc	ccacccccac	ttcactctga	61680
ggttcttttg	gttacttctt	ttttacaaca	aaggtttgat	gtattcttta	aaataataaa	61740
agaaaaagcc	actatattgc	ctcagaacta	ctccatctca	ctgcctcact	ttaaattctg	61800
attgactgta	atgagaccac	agacagattg	atgaaataaa	tttagatact	aaatagcatt	61860
gtacttgggg	tgatttagtt	atatagctca	tttatttctt	caacaaattt	gttgaacagc	61920
agtatttagtt	aaggcaatgc	aagctgatgt	cataaataaa	ccccaaaatc	acaatggcct	61980
aaaaacctaa	aggtgcaatt	tttattcatg	gaagcagggc	aatgtaagta	actacaggat	62040
tcaggcagcc	agaattcttt	catcctgtgt	ttccatcatc	ttctagagct	tcaaagtctt	62100
ctgcatttat	ccaagataag	agaaaagaaa	tttgagaagg	tgtacctgtt	tttagccatc	62160
ttctgtcac	ttgtcttg	aattaacaaa	tcacacaacc	atacctaagt	gcaaagtgg	62220
tgagaaatgt	atttctctgat	gggacaacca	ccttccagtg	ataacttcac	actctgaaag	62280
ggctcagctg	caacacctgt	aataatactg	tgtctggacat	tgtggaattt	tctaaaagaa	62340
acaaaactgg	gacctgtgt	tccacgagca	tgctcagttag	caattttttg	caaggaattt	62400
attcccttac	caataataat	agttaatata	tatataaagt	ttgttatgta	tcaggctcat	62460
tattggacca	gaggctgtgt	gggtgggtatt	tcaggactca	aaccctggta	tctgcgctgt	62520
ccaaattcat	gctcttgctt	cctgccctcc	gtagacttct	aatcacttaa	ggcaagagac	62580

tcaagtctag	ttctcttaga	agttcctacc	acagggcaga	gcccacacca	gacagtgaag	62640
gaatgtggat	tatactaaac	aaaaccctag	gatcgattgt	ataggttaact	gttggcaata	62700
aattagagct	caattccaac	tcataaaaag	ctgtactcca	aaatacaatt	caccataaat	62760
gaaagctttt	taagatatg	agttgcctat	actataaata	ttgatttttt	ttcatctgag	62820
tcaattgaga	tgcaaatgct	aatttgagta	gatgagcata	tgagcattag	tggaaaaact	62880
cacctaaaaa	ggtacaagag	agagcaagat	gcccagata	gattaacaca	acactttaac	62940
cagaatgttc	tatcattttc	tttctgcaca	tacaccttct	aacttgcaaa	tctgtcttgc	63000
ttcacatgat	ctcttcccat	ctaatttgat	gggtttttct	ctgaaagaaa	tgaagggagg	63060
attcctccag	ttcttgtttt	ctatggtaat	tatcctaagt	tttaggggaa	gaattttaaa	63120
agaggaattc	taggcacttc	tcaacttgga	gcagaataga	gatgtattcc	agtaagggaat	63180
atagttcattg	ggcataaagt	acttgaatat	gaggtagaat	tgggggtaaa	taacctatga	63240
aatgtataga	tgatattcta	tcagacaaga	gatattacct	atagttttac	aaaagaaata	63300
aacttttagc	taggcctttt	agaaacatgt	aaaattaact	gagtttacca	agtatagaaa	63360
aaataggaaa	ttgtcattta	gaagggcatg	atctattcag	ctagccaacc	aagaaatatg	63420
tcgagtgtc	cgatgcaagg	tagggatga	agagcagtgg	ggagaagggg	gcagtgaana	63480
gtgcaagaca	acacctattc	ccatataggt	cagcagctaa	gagcaacaac	aaaagagaac	63540
aatgggcatg	tacagtgcac	ccacctgtac	caggtaacat	gccaggtgaa	ttattcaaat	63600
tatctcattt	ggcactaatg	catgtaaaag	acatagcaag	gtgcctgtca	cagaaagaat	63660
gtacaataag	tgatagctgt	gatttaaatg	ccagttttat	gtggagggtta	ttatctgttt	63720
tcaaattaaa	gaaagtgagg	catggtgagg	ctatgtgatg	ctataaactg	aatgttttta	63780
tccccacaaa	atttatatgt	tggaaacctaa	tcacgaaagt	aatcgtgagg	ccccgtaaag	63840
gtaggggtctt	tgggggttga	ttaggtcatg	aaaactccac	cttcatgaat	gctgttataa	63900
taataataat	tagtactctt	ataagagagg	actgggggag	ctcatttgtc	ccttctacta	63960
tgtgaggatg	tagcaacaaa	gtaccatctt	tgaagcagag	agtgaactct	catcagacat	64020
tgaatctgtc	agtacattga	tcttggaact	cccagcctcc	aggaccgtaa	gcaataaatt	64080
tctgttgttt	ataaattatc	ccatctgagg	tattttctta	cagtaggcta	agatagatga	64140
tatatccaaa	gccccttagc	tggagaaacc	agacatacac	gcatggaaag	ttcaataaca	64200
atgagattct	ttgtgaaagt	tgtctctctg	gcagttctat	tttactagta	gaatattcat	64260
cttagttcat	gggttgttat	tcctcaatct	ggctgctgag	tgccttttgg	gtgtttatga	64320
ttttaactgc	ctgcccctgg	ttctagtgtc	ctgggttagg	caatgagggt	attttgtgca	64380
aattgtgttc	tgaagtgcct	tgtgtcagct	cccctgagct	cttccattct	aattctaata	64440
aacttccagc	tatatctctt	aagcagagta	ccttttagga	cagggtgata	gggggagcct	64500
agaggtcaca	tgcgaggcct	cccatctgtg	gaggtttagt	tggagagagc	tgacaatgag	64560
atgattcctg	tgcagcagta	gtcatgaacc	tggggcagga	atgagagcac	catgctatct	64620
cccagggccc	agagagtgtg	gcctgagag	acagatcagt	gaaatctcta	aagagcaatg	64680
cgccagattg	atgagattat	cttgggggta	cttgctagtc	tctttggctt	cggaggggag	64740
tggagggagg	ggaacagata	ggctcagatc	ctctctgccc	ttgaggctgg	tcaaacatat	64800
tgcaggtttg	tagagaagtg	tcttctcctt	gtactgttac	agatgatttg	ctaaccggac	64860
ctgaaactgt	cctatctccc	tccattccag	tgactcaagt	gtcctcagag	ctaataccagg	64920
taactagagc	caggggcagg	cagttaaaat	cataaacacc	caaaggggaa	ccagcagcca	64980
gattgaggaa	taacaacaac	ccatgcccc	agatgaacat	tctactagta	aaatagaact	65040
tccagaagaa	acagatgaca	actttaacaa	agaatcaaaa	gtatacttac	aaaaattgaa	65100
aaatagcatt	aaaataactt	tccagcattc	aaaaacatgg	cttttcaatt	aaaaacatca	65160
attcagttaa	ggaaaggatg	gcctaggatg	agacagactc	acttgcaata	taagacatta	65220
taggggtaaat	accaaaatgag	tcgttttgat	tctcaggtgg	aattgaccat	tagactcacc	65280
tgtaaaaaac	acagacttca	acttaccag	ccctaataaa	tcaggaaactc	tggagccgga	65340
gccaatgaaa	gtgcattttg	gttggaagtg	gccactgggt	ttgggcactg	gttatggggg	65400
ctgtggaaac	tggcagtaga	tggtgaatga	gttatcaggg	tgagagatca	ctgggtctgc	65460
agggatgggtg	gtaaggaccg	cgggagagag	atgagaactg	aaatgagcat	gagaggtggg	65520
tagagtttat	tgatagtgc	ctaaaaatat	gaaacaagga	ttgatgaatg	ttttgtctgc	65580
atgatcccaa	gttttcacca	gatttagaca	caaataactg	tcacaccgt	agatcaataa	65640
gcaagtctaa	aattttttta	agtttagctt	taataaatat	aatacacatc	agataaataa	65700
gaagttgaag	tatcgcgatt	ttgcaaatgt	ttaggaagct	tgtgttcac	cggttcacat	65760
ttccttcact	ctatgtgtgc	atgtttgaaa	ctgaaatata	caagttaaat	ctttagtctt	65820
acttttttaa	ccactgacaa	tgatgatggt	ccatttatct	cccaatgttc	catttcaaga	65880
cagatggatc	agtttgcata	ggtctgagta	ggaataaatg	gagtgtctca	aaggcaaatc	65940
tgtagtgggtg	tgttcaattt	gataaaatca	cgtctgcata	ttttttctct	aaatttctgc	66000
aacctctcca	ctaagggaga	caacctagta	cttcttggtt	tttttctctt	caaattggctg	66060
cttctgagtc	acagatgtga	aataacagga	atcaaaggat	tcattggatca	agaaaaatga	66120
ctcctgatca	tttccagtag	gttattttct	tggagctgtt	ctgctctgcc	cccaggagggt	66180
gtaattggatt	tgaactaaat	gtaatcaaaa	attttcacct	gcatccttat	ttcctgaaag	66240

tagctcagtc	ccatgaagct	gaactttctg	cagctgtgag	acccacggca	tggctgcttc	66300
cttaggggaa	ataaactcga	tgtttctctg	agtcagaacc	acagaacaac	agtctcatga	66360
aatgctctat	gccaaacaaa	aagcactctg	cactcagatt	tgggatgaga	ttcacatgcc	66420
atcagctctc	agagggtgat	agagcacatc	caagcttctg	gagccctgc	agcagagaaa	66480
tcatttttatt	cataactctc	tgttaacttag	tcacttgtaa	tcatttttga	gtgctgcctt	66540
aacaaagtaa	caactgggta	gttttagacaa	cagaaatgta	ttgtctcaca	gttccagagg	66600
ctagaagttc	aaagtcaaag	tgttggcagc	tctgtcccaa	gcttccctcc	ttggcttgta	66660
gatggctgat	ttcatgttca	cagggcattc	tccccatata	ggtgtctatc	tccaaattcc	66720
atctgtttgc	aaagacacca	gtcatagtgg	attagagtcc	atctaacgac	ttcattttaa	66780
cttgattacc	tctgtaaaga	ccctatctcc	acataaggtc	acattctgag	gtactgaggt	66840
ttagaactta	ccctgaatt	tgaggaggga	gatgtaattc	aactgattat	accactgaat	66900
ccattttttc	ccatgtcatc	tattaacagc	cgttagaaga	cgttttagagg	gatgttactg	66960
tggcttcaca	gataccatgc	tcttctctca	tggttcccaa	atgccatttg	cagaatatat	67020
gcagaatcac	tgggagtgc	ttttagaatt	ctaaatctgc	caggcatggt	ggctaacacc	67080
tgtaatccca	gtactttggg	agaccaaggc	aggcagatcg	cttgagccca	ggagttcgag	67140
accagcctgy	gcaacatggt	gaaaccccat	ccctactaaa	aatacaaaaa	attagctggg	67200
tgttgtggca	tgcacctgta	gtcccagcta	cttgagaagg	ttaggctgca	atgagccatg	67260
agcccatgat	cctccactgc	actccagcct	gggcaagaag	agtgaacctt	gtcaaaaaaa	67320
aaaaagaaga	attatagatt	cttgggtctc	acctccacaa	attaaaaatcc	actggatcta	67380
aggaggaact	caggaatttg	caattttacaa	agttgcccag	aacattctga	tgcataacca	67440
ggttggagtt	ccacagccag	ttagtgtctc	cttattattt	tttagacttg	atttgggcca	67500
taaatcttat	aatagctacc	taagaagtac	acttcttctt	taaacaatat	gaacatagaa	67560
ttgttttctc	atctctgttt	ttattacatc	ttctatcttc	ctccactgca	gtacaaaacca	67620
cttgatgtgc	atagattctg	ccttactcat	ctttttattc	tcagtgtgta	ccagagaatc	67680
cacacataaa	taacattcaa	taagcgttta	ttgaagaaat	aaatacctcc	agcaacacag	67740
gcaaaactga	caattgtttc	atcattttaa	gaagttcctg	gccgggcgca	gtggctcatg	67800
cctgtatccc	agcacttttag	gaggccaaga	caggtggatc	tcttaagctc	aggagttcga	67860
gaccagcctg	gccaacatgg	tgaaacctca	tctctattaa	aaatacaaaa	aaaaaggtag	67920
ctgggcatgg	tgggtgtgcac	ctgtgtgtccc	agctactcag	gtgattgagg	caggagaatt	67980
acttgaacct	gggaggcgaga	ggttgcagtg	agctgagatc	acgtcactgc	actccagcct	68040
gggcaacaga	gtgagactcc	atctaaaaaa	aaaaaaaaga	aaaagaaaaa	aagaaaaaga	68100
aaaaagaaga	aaaatgttat	tgacattgac	atatctttat	acttgagaga	attcagctac	68160
tcttaagagt	ttaaccctgt	cagccatgcc	cgctgtcctg	actcattttt	actcctgcca	68220
tgatagctta	cctgtgggtt	caggacacat	ctttcttttt	ccagcctcaa	ggatcttttg	68280
cacatgctat	tttggaaactc	ttctgcctac	cctctttctc	caccccatcc	ccagttgcca	68340
cgactatttc	caaatagaata	tttgttctgc	ccttcaaatt	agaagtgggt	cccctgttaa	68400
acgttttcat	aatgcttcat	tttttttccc	atggagacac	tgatcatgct	tgtaatgaat	68460
agttcaatgg	tgtctctctc	tgacaaactc	taatctacat	gtaggcaggc	ttcatgtctg	68520
ttgtgtttat	ttttgaagcc	acagaaaact	gaacagagta	ggaactcaaa	tattaattga	68580
gtcgttatgc	gctttaccat	gattatgtgt	tttttccac	cattaaactg	ggagcttctc	68640
tacagcaggg	actagattta	atttattttt	ctatttttgg	cacctggcac	agtgcctgga	68700
aggcagaaca	agcacaataa	atcttcttaa	aatttgtaag	gtttttgttt	atttgaatgg	68760
ttgttttagtt	tttgccttga	tgaagcctg	ggatagaaat	ttagaacaag	gttggatggt	68820
gggtgatcata	taaaaataga	tttgttcttt	tttgtcatta	gagcccaaac	ttaaaaaata	68880
tgtcaaataa	aaaaataaac	tagaataaca	gggaacattt	aatggaacca	aggcacctca	68940
aacagctgac	atacttctaa	aaagcagaaa	gtaatatgta	gggagataat	agcatacgag	69000
ttagttttcc	aattgccaac	agatcacttc	aaatacagct	cttggcagtc	atcttcttat	69060
ggagaaaactg	caacacttat	ggaatggcaa	tccaccaat	cccctttttc	taacagcaga	69120
aaaacctgaa	ggagacaagc	ccagctatag	tccagctggg	acctgtggca	ataacctgac	69180
ctaaagttag	ataagtgtac	aggaccgtat	gtgggtccgta	cccctctgcc	tgaggatgct	69240
acgcagctgt	ttactctctc	caaacatttt	gttctttttc	tgacaaccca	ctgcagtcaa	69300
gttgctctatt	cccataatat	ggaatatgtg	gtctatagta	atttcctatt	cctatttcta	69360
ggcaatcatg	taagataagg	ggaatgggca	cctatttctc	tggggaagggt	ataaatgtgt	69420
cttttatctc	tttccagtat	ccccactgtg	tagagctggg	ccacaatgga	gaggaggaaa	69480
agttttgcat	tggaccaaat	cactgcctta	gaaggaaaaga	gattgcaatc	ccattcattt	69540
ttaagatctg	caatgggttg	ggatgatacc	catttaaaaa	agaaaaaaaa	agaggttacc	69600
tctttggcct	ctttggcaaa	ggatatgaac	tagtaataaa	gcctcagcat	caaggcttgg	69660
gagttctggct	caaagtggct	aagactagaa	gcactttcta	acataaatat	ggcagcatta	69720
ggaaaaatagc	ttggccttta	gagagaccca	gattatcttc	accttaaaaa	ctaccaagac	69780
cctggtccag	caagaccacc	atcacagcag	taaatacata	gccaccaact	tcacagaagg	69840
ttttccactg	agacattgtg	aaaaccacaa	tccattaccg	gtggatgagt	cccagctctc	69900

ttactacctg	gagcttccaa	gcaagactta	cctgtccttt	gttccagatt	ggttcctggg	69960
atcctggacc	aacacaaaga	gaagaggaaa	gcctcaagg	aaataaaaagt	gaaggtttca	70020
gtgatagttt	aaatatatat	ttttaaatgt	tagctttttt	ttaaatactt	cacagtgatg	70080
gatggagtat	ttcaaccaca	atggcaaaga	atgaattacc	ggcagtgtca	gagtttcttg	70140
ttgctgtggg	aagacagctg	ccatgttgag	agggtgcctg	tggagaggct	catgtggcga	70200
ggaactgaga	ctggcttcta	tccaacaacc	agtgaggaa	taaggcccta	aagtcctata	70260
gtccacaagg	aactaaatcc	taccaacaac	tacatgagt	agcctggaat	ctgattttcc	70320
ccagtgggga	cttaaggtac	ctacaacctg	gctgttatgg	gctaaactgc	atcctgcca	70380
aattttatatg	tttgagtctt	aactcccagc	acttcagaac	atggttttgt	ttggagacat	70440
agtcattaaa	gaggtaatga	agttaaaatg	aggtcattag	gatgagtcct	aatccaatag	70500
gactggtgtt	cctatatgaa	gaggaaattt	gaacacagac	acataacag	gaaagaccat	70560
gcgaagaaga	ccagagagg	aaacggccat	ctacaatcca	aggagagagg	cctcagaaca	70620
agccaaccct	gcagatacct	tgatctaggg	atccagaatt	gtatgaaaat	acatttttgt	70680
tattttaagct	acccagtctg	tggtgctttg	ttatggcaat	cccagaaaac	taattaacta	70740
attgactgca	gccttgcaag	ggaccctgaa	ccagaggacc	cagttaagcc	atgtgcagat	70800
tcctgactca	cagaactagg	agataacaaa	tgtttggtgt	tctaagggtg	taaagtgttg	70860
gaaaatatgt	tatgtggcta	tggtgttttg	cacatgtgcc	acataatgat	gttttggtca	70920
acaacagacc	acacatatga	tggtgggccc	ataagattgt	aatactgtgt	ttctattatg	70980
cattctctat	ttatagatac	acaaatactt	gccactgtgt	tacaattgcc	tactgtcttc	71040
agtactgtaa	catgctgtac	aggtttgtag	cccagaagca	atagactata	ccatacagtc	71100
tggtgtgtga	gtaggctata	ccatctatgt	ttgtgaagta	cactctatga	tggtcacata	71160
attatgaagt	tgccatagtga	tgcatctctc	agaatgcatg	cccagacatta	agtgacactt	71220
ggctttatgt	ctatcatgca	tttaagatat	tgcaatatag	tttaacttatg	ttttaataca	71280
cctcttaatt	ctgttttgat	cccacaccaa	ccaagatgca	aaactaattt	tttcgcccac	71340
ttccataaaa	aacacaatta	caatgggcct	aatgtggtt	atgtcatttt	ccctgaagat	71400
ttgtggtggt	ataggaagca	gtcacagcac	cctgtataaa	gctcttccat	ggcctgccc	71460
cttttcgcag	tcacctctat	tatttgaagg	aatggtccaa	ccaacctcaa	agctctggaa	71520
atltgcccac	aaataatttt	accctgaagc	cattgaagct	taagtttcag	acccagttt	71580
cacattcctg	tgagttctgc	aagtcactgg	gagctataga	gtcttctacg	tagatttggg	71640
aaacaagggt	aaaatacctc	aagaaatctc	agaagccagg	ggcctgcac	tccaagcctc	71700
tggaacattt	cttgtggttt	ctatcctcac	tctagatat	tgcttttgag	actgattttg	71760
aattcagaat	tttatattct	ctcaaaaaga	atattctcca	aattctataa	gcttcaggct	71820
ctacaaaact	tgtctattct	gcctaactag	atttatcata	gcacgtacca	aagatggccc	71880
tgccatcatga	ggacttttca	tatcagaccc	ttagataata	atattccctc	atgtaagtaa	71940
taatatttcc	gcatgtagtt	atgtatgaat	aattttcatt	catctatatc	atgttgcca	72000
tagtcagtg	ccatgtccaa	tttatttttt	gtgtccttgg	tgctcagaaa	gttaccagat	72060
gttcagtcag	aatttttatt	aaaatatatt	gatgaactag	atlttaaagat	agattaaatt	72120
aattccatac	catgatttaa	tttaattgat	aaaataagg	actgagcatt	agcctgaact	72180
gtgtacttgc	attagccctg	atgttcccca	gagggataaa	tatgtgaaga	atactgcata	72240
ggctgaaaaa	caagagttcc	cagcagacca	atattctttc	tttctgcaag	ctgtgagaca	72300
cctggggaaa	gttctccaca	tttggttttt	cagctactcc	acccatgact	agaaggaaga	72360
ttttatttta	ttattttatt	ttattgtatt	ttttaacagt	aaaagaaaaa	ttattgttat	72420
tttaatgaaa	aagaaccacc	tggaagtct	agtgacattt	ttgaaggcag	ggaccacttt	72480
ttccagggat	ttgtttccct	atgctagtgc	tccattcacc	cttgaccac	ttcagcaaa	72540
cacatagtag	gtgtagcagg	gagatgttaa	gatggaaaact	tgccctgac	cataggccaa	72600
gccagggtgt	ctagtgcgga	tttatccata	aacatcaata	tttatcacta	gatgatatat	72660
tgattttctta	tctgcctcct	atacctattg	atctattttt	cacagcttat	agtggggagg	72720
gtcatgttca	gttattatcc	ttaaaccaa	acaatctctt	cttgaaaaat	aggatgctat	72780
gctctgatgt	ttgcagaggc	caggaaattc	ttgctctgag	agttccagaa	aaaaagattt	72840
gtccagattt	gaactaattt	cagaaaaaca	aaacaaaaac	aaaaaacatt	ccaaggttaa	72900
gaggtgactt	ctgactcaca	gaatgaacaa	gaaggcagg	gaaaaaataa	gtccccaag	72960
attcaggtca	aaggaaatga	acagaaggat	gatggagtgg	cccagaggag	gatcacgtga	73020
cagaaaagtc	aagataaatt	agtggcagga	cagaggaaaa	gaactaaaga	cctgcaggat	73080
cctgaaagaa	agaatgaaag	gaagcagatg	aagattctag	agcccagctg	aatagccagg	73140
ccgcctgagg	gtaaagagg	agagaaagag	gaagcagtga	atcaaatcaa	aaaaggatag	73200
aaaaatacac	agaaacaagc	cagggcacgg	tggtctcac	ctgtaatccc	agcacttttg	73260
gatgacaagg	aaggcgatc	acatgaggcc	aggaatttga	gaccacgctg	gccaacatgg	73320
caaaaccctg	tctctactag	aaaaatatgg	aggaattgtt	gagagattgt	gcactgccag	73380
gggtgaggac	ataaatttat	ttgctatgtg	aatgtttttt	gagactcatc	tagttccttt	73440
acaattttta	tttagaattt	ctttagctcc	aatcaggat	cattagctgg	gagttcttat	73500
aatcaaaaag	caaaacaaa	ctattgcaaa	caagtgggag	ttctcataac	cttaatcaga	73560

aaggcttgag	ttgcatttcc	aaatcctgtc	tctagagacc	atgagttttt	aatgatagta	73620
tttccatgaa	agcccaattg	aatacaatca	ccaccagac	caaagtgtag	gggaaatcat	73680
aaagctactt	tcaaattccga	gtcttctttg	aaatttaaaa	ttatattcca	tgataatggg	73740
tctcataaac	cttaattgcaa	gggaaaaaaa	aaaaactaca	gcagaataaa	gaagatttaa	73800
aaagaagcag	gataggaaaa	aagacaagcc	agaggctagc	tatatctttt	cataaagcac	73860
acctggtctc	ataagcacat	gagagctagg	cccagaagtt	tcgagccttg	gagactcact	73920
tcaaacaaat	tccagggaaa	gacactcaat	taagtgttca	ggacaattgt	atatccccag	73980
tgcactcatg	tactggaagt	gcttaaactc	tgccctcagt	ggtcaataaa	atggtattaa	74040
gtcacagaca	accatttgcc	attggtattg	gctgatgtct	ctggtagttt	atgtgatgac	74100
aaacattagc	cttaacagca	gagccttcct	ttctttgatt	tccgtgtgaa	attctgtggc	74160
tgggagatgt	ttaggttcac	atgaacataa	ctgtaagcaa	accatcttat	ggtagacatc	74220
tctccagaaa	aagcaaatc	catcttttta	aatagaaaaa	aaaaacagca	gcaaaatgaa	74280
attaatgcct	ttggctcatt	gtgttgtaaa	attatccttg	aagttatttc	acactgccct	74340
aaaaaatcat	tgaattcaga	ttataacaca	gctcttatga	agctgttagg	cagttgcac	74400
tttaatgagg	ctgagtaggt	aagaggaaca	gatactctag	attgaaagcc	atcgcaaaaa	74460
atgtgaaaca	actgtattaa	atttaattta	atttaaccaa	tgggtgtatt	tattccctgc	74520
cttcttccag	aaaggattta	gggagcttag	gtaattgtga	ggtctgaagg	tgtagatact	74580
catgaaagcc	atctaataat	atctagaaca	tataattgtg	ccagatagtt	ttcaattatc	74640
aatgcattaa	tctaagcttg	gaaaggactt	ttaaactgca	atgaaacctc	aactagtttc	74700
aaacctgcat	tttttagtaa	gctattttgt	catttgga	cccacacaaa	aaacctattg	74760
ttcctttaaa	aaggagagag	agaaaaaag	aaaaacaacc	tgctggtgtt	gctttgttca	74820
tcttatcacc	ttatcattca	gcagcctatt	caggatcaaa	tgtgggaaag	tctgtgtgtg	74880
agattaagcc	ctatactcag	cacaggaatt	tcttgataaa	atgtccttct	taaatggata	74940
gctatgtaaa	aaacagaaga	gaaactctat	tattacaatg	cagttgggtt	tttttgtttg	75000
gtatgtttca	aatctaatat	ttaaaattaa	gaaaattcaa	agcaaaccat	ctgtgtggat	75060
gttttcatac	tccatttttg	tgacaatagc	ttgtcttaacc	gagattatta	gtggattcct	75120
tcaataagag	gaagagagta	ccatcattaa	gatggtatgc	ctccactagc	ctttattttt	75180
gagattaag	aatgacaaaa	atattttatat	atattttttc	attattaaga	cggagtctca	75240
ctctatcgcc	aaggctggag	tgcatgtggc	ccatcttggc	tcactgcaac	ctctgcctcc	75300
caggttcaag	caattctccc	acttcggcct	cctgagtagc	tgggatcaca	ggcaccgcc	75360
accacacctg	gctaattttt	gtatttttag	tagagatggg	gtttcaccac	attgccagg	75420
ctggtctcaa	attcctgacc	tcaagtgtac	cgctgectc	agcatccac	agtgtgtgga	75480
ttacaggcat	gagccaccac	aactagccaa	aaaatatata	ttttctttaa	aaaaaagtta	75540
ttttctgcct	aattacttca	atatttcatg	aacattgtgg	gggaaagaat	attcatacag	75600
ttctctaata	atatcagtga	gttgaggtac	caaatacaga	attctattga	attctaattg	75660
caacactgaa	gcagacttac	tctaagatcc	tgctacctgg	actaaccaca	gactataccc	75720
agattccatc	tgcaagatgt	agtggagcca	aacatcacat	cactgaggaa	tatttattat	75780
atatctttct	gtagaggtga	gagagtattt	atagatataa	caaacaattg	gtcagatact	75840
actggagatt	tgggaagctt	actaggtttt	tcactaaatt	tgttgatagc	caactctaaa	75900
gctcattcta	ctactactag	ccatcttccc	ttggacaatt	tcattgggcat	cgcttactg	75960
aaaagaaaaa	tgagtaagta	taatctcact	agtacatgaa	taactattat	aaacacaata	76020
ctttctacaa	ggtttattac	taatctatgt	catttaagtt	ctagatttta	cagtattttg	76080
gagtgatgaa	taaatagcat	gctgttcttc	tatagctagc	aaacaagact	cagctttgga	76140
gacaagaaga	aggaatacag	gagatgtttc	cttgaatgag	tcttcttctc	aaactctgct	76200
agaagtatgt	cagaaagcga	aggttacctc	agtctccaag	tcattgaaaac	atgaggtctt	76260
tctgttttct	agggaaagta	tttcatcatt	tcctagtgtg	gctttttgca	cacccagaa	76320
ggagggaagac	aagccatcgt	gtcaccttta	catggtagtt	gagatgcttc	acatttgtac	76380
ataacatctt	atatgatcga	aagaggaact	caaactcgac	aggggagaat	aaggaaatac	76440
atctaaaact	gacaattatg	ttaatccctc	aattccctgg	gccccaaaaa	tagtctcctc	76500
tatttgccat	ttcaaaatag	ctaataagag	ggtaattgga	catgtagaca	ctgaatttcc	76560
actgctcatg	gtttccttcc	ctgggctccc	agtcacccgg	aaggccccgg	attacctcaa	76620
atgtcataac	gggttcacg	cattacgaat	tacctccttc	tgccaataaa	ctacaagttc	76680
tttgaggaca	gaaactgtga	tttatgggca	tttgatccc	cagagattca	aagagtgcct	76740
gatacagagt	ggtgatagat	gattttatca	tcttatcatc	atcatcctca	tcactgttgt	76800
acctaaatag	aggttacttc	ctgtttcttt	tatttaacaa	taacctatta	taatataac	76860
tatgcatttt	taatgcactt	tacaaatatt	gtcctcattt	aatccctgta	aggtgaatac	76920
tattattggc	tccattttgc	agataagaca	ctgagtcgtg	aaaaggctaa	gtacctcaca	76980
caagatcaca	cagccagtg	cagaactagg	atttgctgg	actcttaaag	actctggttt	77040
gccattttaa	ggtacgctga	gctccacaac	actgccttag	aacttgctca	gagcaattgt	77100
catattacta	tggtaggttt	ccctgtctcc	cttaagactt	aaagcttctg	gagatcagaa	77160
actgtattgt	ttacttgcaa	aatcatgggc	ttaggagcca	gaaagatctc	gatctaagtc	77220

```

ctgcctcatc tgccttcttct tcatgggatt agtagactga ttttccctt gcttagccat 77280
gcagtttgaa tgggaattcac cccattctca gctacaacca cggatcttgt ttgctttaaa 77340
ccaatcatca ctgcccattt ttctggcctt tgattgattt aaagggtgaac acatgaccgc 77400
tttggaacaa tgaggctcag gtggtatttt gtgagggtct ctggaataga aaatcatttc 77460
tcttctgaca gaactacaag gcaaggcgca atcactgatg gtgaaaata aagatcatta 77520
tattaggaac tgttggtgtg catactgcta atgtgagggt agttagcctt agaaagagga 77580
tgctaccttg aaaggcaggg cagatctatg gaaataaact gagtttgtgg tgacattgat 77640
agccacttta gttcatcagc tgaatcccaa tctacttggt ttttccagct ctgtaagcca 77700
acaaataccc tgctagtgtt ttagcctatg ccagttgttt tttctgctgc ttgttaataa 77760
gacacatata gtcttggaac aattatataa aaatatgaat ctttgccag gcatggaggc 77820
tcacgcctgt aatcccaaca ctttgggagg ccgaagcgtt cggatcacag ggtcaggagt 77880
tcgagaccag cctggccaat atggtgaaac cccgtctcta ctaaaaatac aaaaattagc 77940
tggagtgtgt cgcacatgcc tgttgtccca gctactcagg aggctgaggc agaagaatcg 78000
cttgaacccg ggaggcagag gttgcagtga gctgagatca tgccactgaa ctccagcctg 78060
ggagacagaa tgacactcgg tctcaaaaaa aaaaaaaat tctgaatctt tgaatgttat 78120
cgtatagctt tttaaaagac tagattaagc tatggatgta agcaccaatg ctttttttta 78180
aaagtatatg ggatgaatga gttaaagtct aagtgaattc agcaagtttg taacccttga 78240
gctcttgtga gctcaatctg cctcaatca tctgagaatg aacaaaacag actcaaccat 78300
tttgtgtcat ccagatgaat agaaagggcc tgaatatatt ctttttttat attttttgtc 78360
ataggcccta gcgttttcat ctgctcagtt tttgagtgtg cttgttaggg ctttaatgct 78420
aaactcaatg taagctgggt ttatgtaatt tagtggtttc tagacttagg cttcctaaat 78480
gttccattag acaagaagac aaaatatgaa acaaaataac atcaaaactac cattaacaat 78540
tagcattgtc agagtacgta aaaactgttt tgaaataaag cagcttttca ggccagtatt 78600
tgagaactaa tgatagagtt tttttgtagg caatgaatga agtgttgtgg aacaaatgtg 78660
gatttggaat ttcaagacca acttgaacca gagtcttgaa tctgcaactt tactagctgt 78720
ggcttctttt taattgatta actttctttt tctgtaaaat ggaataatg attcctacct 78780
cttacaagcg ttgtggtagg aggtattcca aatgagataa atatgtgaaa tcaactaaaac 78840
ttccagaata catggaaata atatatgtct gtattggcaa attatgagtg tagtaatgag 78900
ggatttacat ggacacattt acattcttta gttcagactc ctttgacga cacataacaa 78960
acacacacaa gaatatttgc cctgtgtcca gcactctttt gagaatttct ctcttctcct 79020
ccttctgata tcccaaccaa ccaaccagta ggaatgtgcc ctagaaacta tgttttacag 79080
ctcagttccac atcttttcac cttacaaatt acttattgtt agtttcttgg gaaagtaaaa 79140
ttgagatatt aagattctat gaaattgggt ggtggtattg aagactgtgc tgctgctgct 79200
gctactgtc cagaactcta gtggagatat tacaaatggg agcagaaaaa gtgagaaggt 79260
ggtctgtaga gacagagcac agcacagagg tagaaagggt gagataaaga aacaatgcag 79320
tgagagaagg agagaacaca gacatagaaa cagagacata aaggcaccag acacgagagc 79380
agagataatt taataaggcc gtagtttctc ttggatttcc agttccttct tttagtcaat 79440
tctgatattt agcacactgc actgcaccag atatcttgtc gccttcaa atagattttt 79500
tcccctccac acatgcta at ttgggcttgt attatttgca atcaatagaa ttgagaaaaa 79560
cactctatgt catatctgaa ctagtatatc agaattggtt tctgccattg gtaacagaga 79620
cccagctcaa tggcttaaac acacaagggt ttatcctgtc acattaacaa aaaaaaaagt 79680
gcagagtagc agtccgggggt tgggaatagg tccacaaagt catctggggc caggcttcat 79740
ctatctttct cttacatcac aagtgacttc tgtctacaag gcccctcatg gtccacaatt 79800
gattgttttt acaacagctt tccaaagagc agaaagaaaa acaagtagaa gagcttgtac 79860
tttcttttct atttatgtca cctctgctta catctcacta gccagatata gacatatggc 79920
cccattctacc tgcacagaaa tctcagaaat gtaatttttc agcttgcca ttgccaggg 79980
ttctgttact agggaagaag 80000

```

<210> 6

<211> 2260

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a cDNA detected by cDNA cloning. It includes two exons. It is transcribed from the +strand of Seq_ID_06.doc.

<400> 6

```

atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcacatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaaggga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttacac ctttggtctc cctggttctc agttcttttg acttggactg aattacactg 240
ccaggtttcc tggttctcca gcttgcagat ggcagatcat gggacttctt ggcctccata 300
attgttttca tatctccagg cctttcattg ggtcagggtt gcatttcgct gccctttatg 360
tgtgtgacaa gtgaaaataa ggaaagaaaa aaactcaagt gaagaaaatc agaactctgcg 420
cagcagttcc tgggcgtttc agctgcttcc cacatcacct gcctcatcaa gccccagcat 480
ccatctcctt gctcatctta caccctgtgt gcatgacagg cccaccattc atttatcaga 540
gcaaaggctc tcccactatt ctggttcacc cccctactta gccagatata caagaatc 600
tgcacggatg acctgcctca cctgggagct cagaggagct cagattccat tactatcgca 660
ccaaggacag atctcccagc aagaatgaca gaaaagacta actgccccca aaatctccct 720
tccaaaacac agttctctta attctcccaa gaaaccagaa tgtgactgct cacctctcta 780
aggacctgaa aacaactggc catttcagct atttaaatac actttaaaaa atccaaccgc 840
caaaatatta aaccattttg gttggaatga taacataact aacctgctga cagctgcttc 900
tgctaggtgc aaaaatggaa aaaaaaatc ttctaactag gtcaaatcac tctacctttg 960
ggattctaaa tttactcata ttctcaaaga aatatattca gtcatagtgg ggaaaatagg 1020
attattcctt tagctcgata agcaaccaga agttcttctt tcaaactctg acatttaatc 1080
aatcagaaat tgatttttgg aaaactgttt cctatgaagc tatctctgcc tgaaggattt 1140
ttcttttaca atccagacta tagaaggaaa ttcaaacctt ggactttcac ctccattggt 1200
cagagtttta ctgaccaatt cccacctctg ccttacacct aacggaagtt tatgcctgtt 1260
ttctcttcac ataccccaac agttacaaat ggttgttatt attaagcatc ttttattttg 1320
tggcctctga ttacatggtc ccctaaattt tgacctaatc acaaagattt ggtaaaattt 1380
cttaacatat taataatatt ttgtttatgt gtcaatatct tagcatgtat caattaagac 1440
agaggtctta acgttctctt ttgaaagag aatattagga ttcagagata ttaagagatt 1500
ctcccaggat cacagttagg taacagagct ggattttagt ccagggtctgt ctacagctct 1560
aacgtatata caccctttgt ataacatgtc acgaattcag cataaaggga tcttcagtga 1620
tctaagtcag ggttcagcaa ccttttctaa aaaggaccaa atagtaatat ttcaggcttt 1680
gtggacccta tggctctcat cataactgtt caaatcacca tgtagtgtaa aaggagccat 1740
aagcaaaaata taaactaacg aatgtggctg ttttatggga ttttttttta actctttatt 1800
tacaaaagca ggtggcagat cagaactcac ttatgggcca tagttctctg acccctgacc 1860
tgagaaaatc ttatatttat ggacaacatt tagactgtga cttgccaaagt aagaacaaga 1920
agctctgtca actgaagggtc aaggctggag ttctgaaagc aaagagctgt ctggtgttaa 1980
tgataagtga aatagttaaa gttagaagat cccagttata agaagcacia agaataatga 2040
ccatagactc ctgaacaaga atgtctggac ttctggctta ggcactcttg ttgtatggtc 2100
caggccaagt tacctaactc ctccaggcct ccaatttctt atcattaaat gaagataata 2160
aaagtatttt cctcagagag ctgtaagaat aaactgagct aaccatgtc aagcacatag 2220
aatagggcc agcctatatt aatttatcaa taaatgccag 2260

```

<210> 7

<211> 2817

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a cDNA detected by cDNA cloning. It includes three exons, the first and third are identical to the two exons in Seq_ID_06.doc. It is transcribed from the +strand of Seq_ID_06.doc

<400> 7

```

atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcacatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaaggga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttacac ctttggtctc cctggttctc agttcttttg acttggactg aattacactg 240
ccaggtttcc tggttctcca gcttgcagat ggcagatcat gggacttctt ggcctccata 300
attgtgggca aaaaggaaa agacaaaaca gcatgaaatg atgagaccaa gtgatgaaaa 360

```

```

ttcattcaca atgattgctt tcaagagtaa tttctcttgg gtaattcagc agcctgttac 420
tatggctctc tggagtgata gctaagttaa atgaagcctc taaaagtggg ttatcctgac 480
aagaatatac tcagccaata atgcaacaga aatccattca aagcattcgg gaaaaattca 540
aaagaataaa tattcttttt ttttttttaa agttaatgac ctacgatcca tttcttccct 600
gactaacaag cagcaagcac ttaaaaatat ccagccagga tgaaatagaa acccacctga 660
cttgtaataa tttttgtttg gtcccaggga ctcagattct aagccaaatt ctttgaatga 720
tcttggtcaa tgctcgaat tatttttgcc aacttttctt tatcttggaa aaaaagtttc 780
atgaatgggt gtcaaaattg attagtttta aaaacctttc ttgcagatac gtatggcacc 840
ctaaaactgt attagaaaaa aatttcatat ctccaggcct ttcattgggt cagggttgga 900
tttcgtgcc ctttatgtgt gtgacaagtg aaaataagga aagaaaaaaa ctcaagtga 960
gaaatcaga atctgctcag cagttcctgg gcgtttcagc tgcttccac atcacctgcc 1020
tcatcaagcc ccagcatcca tctccttgct catcttacac cctgtgtgca tgacaggccc 1080
accattcatt tatcagagca aaggctctcc cactattctg gttcaccccc ctacttagcc 1140
agatatataa gaatatctgc acggatgacc tgcctcacct gggagctcag aggagctcag 1200
attccattac tatcgacca aggacagatc tcccagcaag aatgacagaa aagactaact 1260
gccccaaaaa tctcccttcc aaaacacagt tctcttaatt ctccaagaa accagaatgt 1320
gactgctcac ctctctaagg acctgaaaac aactggccat ttcagctatt taaatcaact 1380
ttaaaaaaat caaccgcaa aatattaaac cattttgggt ggaatgataa cataactaac 1440
ctgctgacag ctgcttctgc tagtgcaaaa aatggaaaaa aaaatacttc taatcagggt 1500
aaatcactct acctttggga ttctaaattt actcatattc tcaaagaaat atattcagtc 1560
atagtggtga aaataggatt attcctttag ctcgataagc aaccagaagt tcttcttca 1620
aatcttgaca ttaatacaat cagaaattga tttttgaaa actgtttcct atgaagctat 1680
ctctgctga aggatttttc tttacaatc cagactatag aaggaaattc acaacctgga 1740
ctttcacctc cattggtcag agttttactg accaattccc acctctgcct tacacctaac 1800
ggaagtattt gcctgttttc tcttcacata cccaacagt tacaatgggt tgttattatt 1860
aagcatcttt tattttgtgg cctctgatta catgggtccc taaattttga cctaatacaca 1920
aaagattggg aaaatttctt aacatattaa taatattttg tttatgtgtc aatatcttag 1980
catgtatcaa ttaagacaga ggtcttaacg ttctcttttt gaaagagaat attaggattc 2040
agagatatta agagattctc ccaggatcac agttaggtaa cagagctgga ttttagtcca 2100
ggtctgtcta cagctctaac gtatatacac cttttgtata acatgtcacg aattcagcat 2160
aaagggatct tcagtgatct aagtcagggg tcagcaacct tttctaaaaa ggaccaaata 2220
gtaatttttc aggcctttgt gacctatgg tctctatcat aactgttcaa atcaccatgt 2280
agtgtaaaag gagccataag caaaatataa actaacgaat gtggctgttt tatgggattt 2340
ttttttaact ctttatttac aaaagcaggt ggcagatcag aactcactta tgggccatag 2400
ttctctgacc cctgacctga gaaaatctta tatttatgga caacatttag actgtgactt 2460
gccaaagtaag aacaagaagc tctgtcaact gaaggtcaag gctggagttc tgaaagcaaa 2520
gagctgtctg gtgttaatga taagtgaat agttaaagtt agaagatccc agttataaga 2580
agcacaaga ataagacca tagactcctg aacaagaatg tctggacttc tggcttaggc 2640
actctgtttg tatgggtccag gccaaagttc ctaatctctc caggcctcca ttttcttatt 2700
attaaatgaa gataataaaa gtattttcct cagagagctg taagaataaa ctgagctaac 2760
ccatgtcaag cacatagaat agggccagc ctatattaat ttatcaataa atgccag 2817

```

<210> 8

<211> 1970

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a cDNA detected by cDNA cloning. It includes two exons, the last is also present as "exon-2" in a human cDNA clone (sequence acc. no. BC036936). It is transcribed from the +strand of Seq_ID_06.doc

<400> 8

```

atgtgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcatcatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaaggga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttaaac ctttggtctc cctgggtctc agttcttttg acttggactg aattacactg 240

```

```

ccagggtttcc tggttctcca gcttgccagat ggcagatcat gggacttctt ggcctccata 300
atttgtgtatc cccactgtgt agagctgggt cacaatggag aggaggaaaa gttttgcatt 360
ggaccaaatac actgccttag aaggaaagag attgcaatcc cattcatttt taagatctgc 420
aatgggttgg gatgataccc atttaaaaaa gaaaaaaaaa gaggttacct ctttggcctc 480
tttggcaaaag gatatgaact agtaataaag cctcagcatc aaggcttggg agtctggctc 540
aaagtggcta agactagaag catcttctaa cataaatatg gcagcattag gaaaaatagct 600
tggccttttag agagacccag attatcttca ccttaaaaaa taccaagacc ctgggtccagc 660
aagaccacca tcacagcagt aaatacatag ccaccaactt cacagaaggt tttccactga 720
gacattgtga aaaccacaat ccattaccgg tggatgagtc ccagctctct tactacctgg 780
agcttccaag caagacttac ctgtcctttg ttccagattg gttcctggga tcctggacca 840
acacaaagag aagaggaaaag cctcaaggga aataaaagtg aaggtttcag tgatagttaa 900
aatatatatt tttaaatgtt agcttttttt taaatacttc acagtgatgg atggagtatt 960
tcaaccacaa tggcaaaagaa tgaattaccg gcagtgtcag agtttcttgt tgctgtggga 1020
agacagctgc catgttgaga ggtgccttgt ggagaggctc atgtggcgag gaactgagac 1080
tggcttctat ccaacaacca gtgaggaact aaggccctaa agtcctatag tccacaagga 1140
actaaatcct accaacaact acatgagtga gcctggaatc tgattttccc cagtggggac 1200
ttaagggtacc tacaacctgg ctgttatggg ctaaactgca tcctgccccaa atttatatgt 1260
ttgactctta actcccagca cttcagaaca tgggttttgt tggagacata gtcattaaag 1320
aggtaatgaa gttaaaatga ggtcattagg atgagtccta atccaatagg actggtgttc 1380
ttatatgaag aggaaatttg aacacagaca catatacagg aaagaccatg cgaagaagac 1440
acagagagga aacggccatc tacaatccaa ggagagaggc ctcaagaaca gccaacctg 1500
cagatacctt gatctaggca tccagaattg tatgaaaata ctttttgtt atttaagcta 1560
cccagtctgt ggtgctttgt tatggcaatc ccagaaaact aattaactaa ttgactgcag 1620
ccttgcaagg gaccctgaac cagaggacc agttaagcca tgtgcagatt cctgactcac 1680
agaactagga gataacaaat gtttgttgtt ctaagggtgt aaatgttggg aaaatatgtt 1740
atgtggctat gttgttttgc acatgtgcc cataatgatg ttttggtaaa caacagacca 1800
cacatatgat ggtgggcccc taagattgta atactgtgtt tctattatgc attctctatt 1860
tatagataca caaatacttg ccactgtgtt acaattgcct actgtcttca gtactgtaac 1920
atgtgttaca ggtttgtagc ccagaagcaa tagactatac catacagtca 1970

```

<210> 9

<211> 1989

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a human cDNA sequence (sequence acc. no. BC036936). It is transcribed from the +strand of Seq_ID_06.doc

<400> 9

```

tttgaggcc cggcacagtc tcagttcccc acagcacaga gaggaagcaa ggtctttgca 60
tgccctgttc acttggccaa acattgccct agatgtgcaa ttcaacccat ataaagcaag 120
cgtaattcaa gagccctgac tctttcacca tttcatgatg gcttattact cgcacacggc 180
attgtcaggt tccagcaact ctgtgtgcag caggcagaaa tgcagtgttt atctctgaca 240
ggcaaacatt cttaggctca tcggcatgag tatccacgct gaactggcta tcggggaagt 300
agatccctgc atgggaaagt caaggatatc ccactgtgta gagctgggtcc acaatggaga 360
ggaggaaaaa ttttgcattg gaccaaatac ctgccttaga aggaaagaga ttgcaatccc 420
attcattttt aagatctgca atgggttggg atgataccca tttaaaaaag aaaaaaaaaa 480
aggttacctc tttggcctct ttggcaagg atatgaacta gtaataaagc ctcagcatca 540
aggcttggga gtctggctca aagtggctaa gactagaagc atcttctaac ataaatatgg 600
cagcattagg aaaatagctt ggccttttaga gagaccaga ttatcttcac cttaaaaact 660
accaagaccc tgggtccagca agaccacat cacagcagta aatacatagc caccaacttc 720
acagaagggt ttccactgag acattgtgaa aaccacaatc cattaccggt ggatgagtc 780
cagctctctt actactgga gcttccaagc aagacttacc tgtcctttgt tccagattgg 840
ttctgggat cctggacca cacaagaga agaggaaagc ctcaagggaa ataaaagtga 900
aggtttcagt gatagttaa atatatattt ttaaattgta gctttttttt aaatacttca 960
cagtgtgga tggagtattt caaccacaat ggcaaagaat gaattaccgg cagtgtcaga 1020
gttctctgtt gctgtgggaa gacagctgcc atgttgagag gtgccttgtg gagaggctca 1080

```

```

tgtggcgagg aactgagact ggcttctatc caacaaccag tgaggaacta aggccctaaa 1140
gtcctatagt ccacaaggaa ctaaactcta ccaacaacta catgagttag cctggaatct 1200
gattttcccc agtgggggact taaggtacct acaacctggc tggtatgggc taaactgcat 1260
cctgcccataa tttatatgtt tgagtcttaa ctcccagcac ttcagaacat ggttttgttt 1320
ggagacatag tcattaaaga ggtaatgaag ttaaaatgag gtcattagga tgagtcctaa 1380
tccaatagga ctggtgttct tatatgaaga ggaaatttga acacagacac atatacagga 1440
aagaccatgc gaagaagaca cagagaggaa acggccatct acaatccaag gagagaggcc 1500
tcagaacaag ccaaccctgc agataccttg atctaggcat ccagaattgt atgaaaatac 1560
atTTTTgtta tttAagctac ccagtctgtg gtgctttgtt atggcaatcc cagaaaacta 1620
attaactaat tgactgcagc cttgcaaggg accctgaacc agaggaccca gttaagccat 1680
gtgcagattc ctgactcaca gaactaggag ataacaatg tttgttggtc taagggtgcta 1740
aatgttggga aaatatgtta tgtggctatg ttgttttgca catgtgccac ataattgatgt 1800
tttgggtcaac aacagaccac acatatgatg gtgggcccac aagattgtaa tactgtgttt 1860
ctattatgca ttctctattt atagatacac aaatacttgc cactgtgtta caattgccta 1920
ctgtcttcag tactgtaaca tgctgtacag gttttagtagc cagaagcaat agactatacc 1980
atacagtca                                     1989

```

<210> 10

<211> 3931

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a cDNA detected partly by cDNA cloning, partly by PCR analysis. It includes two exons, the last is identical to the last exon in Seq_ID_06 and Seq_ID_07.doc. It is transcribed from the +strand of Seq_ID_06.doc

<400> 10

```

gaaggtaaaa aaaagaggga ggtctgagaa atagaaatat cagaggaagg aaataaagga 60
gggtgagagt aaattctctt ttagcattca gattccacag attccacaaa tcacatttct 120
ttttttacca actaaggaaa aataacactt gacctaacat ttcattgcag ttagctaaag 180
gatgctagaa aaactatgtt gcagtgggtt gctctaattt cttcaggaat agagaaaaagt 240
gacaaaaaga tcagagaaga gaagaaagga aactatcaga aaaatacaga attggagtag 300
gatataacat atttgggttg aaggtaaaat tttatatgtt aatcttaagt atcttgctac 360
ttcagtttgg tccctggaac agcagcatca gaatctgccg agggcttgtt aaaaaggcag 420
aatctcaggt cccatcccag actcactgaa tcagaatata aatactgaca agatgccccg 480
ggattcatat gcacagtaga gctggcgaag ttccattgta gcctgtgatt gttttctgca 540
acttagtatt tctgagtttt cccaagggaag aaaaccagg ccttagcttc tggcagactt 600
gtgtttctcc tttacttact agctgcatga ctcatgagca aggaaatcaa actttatgtg 660
cctgagtttc ctcatctata aaatggagac tataataatc atctcctagg cttgttttga 720
ggatgttcaa caaatgctcc tttcattcct ctatttacag acctgccgca gacaattctg 780
ctagcagcct ttgtgctatt atctgttttc taaacttagt aattgagtgt gatctggaga 840
ctaactctga aataaataag ctgattatth atttattttc tcaaaacaac agaatacgat 900
ttagcaaat acttcttaag atattattht acatttctat attctcctac cctgagttga 960
tgtgtgagca atatgtcact ttcataaagc caggtatata ttatggacag gtaagtaaaa 1020
aacatattat ttattctacg tttttgtcca aaaattttta atttcaactg ttgcgctgtg 1080
gttggttaag taaaacaaac tcagtacagt agtattcagt acagtattta agcccctgta 1140
cttaaacata ttcctcgtac caatgaagtt acatgaaaag caaatttgtg tgagatatcg 1200
tagatggaag taaattagtc tttatgttcc ccacaaattg aaatgcattt caaaaactct 1260
gtgtgtgtat gtgtgtgtgt gacagagtgt gtgtgagaga gagacagaga gatacgcttt 1320
ggttgcctcc ataagctggc tgctatgatt aataagacca agttttctaa agaaaatgag 1380
atcataacaa aagccctctt tatgactatc ttttatcagg ggcaaaaagg aaagagacaa 1440
aacagctata aatgatgaga ccaagtgatg aaaattcatt cacaatgatt gctttcaaga 1500
gtaatttctc ttgggttaatt cagcagcctg ttactatggc tctctggagt gatagcta 1560
gtaaatgaag cctctaaaag tggattatcc tgacaagaat atactcagcc aataatgcaa 1620
cagaaatcca ttcaaagcat tcgggaaaaa ttcaaaagaa taaatattct tttttttttt 1680
ttaaagttaa tgacctacga tccatttctt ccctgactaa caagcagcaa gcacttaaaa 1740

```

```

atatccagcc aggatgaaat agaaaccac ctgacttggt aatatTTTTg tttggTccca 1800
gggactcaga ttctaagcca aattctttga atgatcttgg caaatgtctc gaattatttt 1860
tgccaaacttt tctttatctt ggaaaaaaag tttcatgaat ggggtgtcaaa attgattagt 1920
tttaaaaacc tttcttgCag atacgtatgg caccctaaaa ctgtattaga aaaaaatttc 1980
atatctccag gcctttcatt gggtcagggt ggcatTTcgc tgccctttat gtgtgtgaca 2040
agtgaataa aggaagaaa aaaactcaag tgaagaaaat cagaatctgc gcagcagttc 2100
ctgggcgttt cagctgcttc ccacatcacc tgcctcatca agccccagca tccatctcct 2160
tgctcatctt acaccctgtg tgcatgacag gccaccatt catttatcag agcaaaggct 2220
ctccactat tctggttcac cccctactt agccagatat acaagaatat ctgcacggat 2280
gacctgcctc acctgggagc tcagaggagc tcagattcca ttactatcgc accaaggaca 2340
gatctcccag caagaatgac agaaaagact aactgcccc aaaatctccc ttccaaaaca 2400
cagttctctt aattctccca agaaaccaga atgtgactgc tcacctctct aaggacctga 2460
aaacaactgg ccatttcagc tatttaaate aactttaaaa aatccaaccg ccaaaatatt 2520
aaaccatttt ggttggaatg ataacataac taacctgctg acagctgctt ctgctagggtg 2580
caaaaatgga aaaaaaata cttctaatac ggTcaaatca ctctacctt gggattctaa 2640
atttactcat atttcaaaag aaatatattc agtcatagtg gggaaaatag gattattcct 2700
ttagctcgat aagcaaccag aagttcttcc ttcaaatctt gacatttaat caatcagaaa 2760
ttgatttttg gaaaactgtt tcctatgaag ctatctctgc ctgaaggatt tttcttttac 2820
aatccagact atagaaggaa attcacaacc tggactttca cctccattgg tcagagtttt 2880
actgaccaat tcccacctct gccttacacc taacggaagt ttatgcctgt tttctcttca 2940
catacccaa cagttacaaa tggttgttat tattaagcat cttttatttt gtggcctctg 3000
attacatggc cccctaaatt ttgacctaat cacaaaagat tggtaaaatt tcttaacata 3060
ttaataatat tttgtttatg tgtcaatate ttagcatgta tcaattaaga cagagggtctt 3120
aacgttctct ttttgaaaga gaattattag attcagagat attaagagat tctcccagga 3180
tcacagttag gtaacagagc tggattttag tccaggctctg tctacagctc taacgtatat 3240
acaccctttg tataacatgt cacgaattca gcataaagg atcttcagtg atctaagtca 3300
ggggtcagca accttttcta aaaaggacca aatagtaata tttcaggctt tgtggaccct 3360
atggtctcta tcataactgt tcaaatcacc atgtagtgtg aaaggagcca taagcaaaat 3420
ataaactaac gaatgtggct gttttatggg attttttttt aactctttat ttacaaaagc 3480
aggtggcaga tcagaactca cttatgggcc atagttctct gaccctgac ctgagaaaat 3540
cttatattta tggacaacat ttagactgtg acttgccaag taagaacaag aagctctgtc 3600
aactgaaggc caaggctgga gttctgaaag caaagagctg tctggtgtta atgataagt 3660
aaatagttaa agttagaaga tcccagttat aagaagcaca aagaataatg accatagact 3720
cctgaacaag aatgtctgga cttctggctt aggcactctt gttgtatggt ccaggccaag 3780
ttacctaatc tctccaggcc tccattttct tatcattaaa tgaagataat aaaagtattt 3840
tcctcagaga gctgtaagaa taaactgagc taacccatgt caagcacata gaatagggcc 3900
cagcctatat taatttatca ataaatgcc a g 3931

```

<210> 11

<211> 9458

<212> RNA

<213> Homo sapiens

<220>

<221> mRNA

<222> (0)...(0)

<223> Corresponds to a cDNA detected by cDNA cloning. It includes one exon. The sequence includes the last two exons in Seq_ID_07.doc. It is transcribed from the +strand of Seq_ID_06.doc

<400> 11

```

atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcacatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaaggga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tctgcccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttacac ctttggtctc cctggttctc agttcttttg acttggaactg aattacactg 240
ccaggtttcc tggttctcca gcttgagat ggcagatcat gggacttctt ggccctcata 300
attgtgtgag tcaatttcca ttttatttac atatccagtt atgcattgct taacaatgga 360
gacaggttct gagaaatgca ttgttaagt atttcatcat tgtgcaaaca tcatagagt 420
taactacaca aacctggaca gcatagacta ctacacatct aggtacatg gtgtagcttg 480
taacctcatg ataagtatgt ataacatcat gataagtatg tatgtatcta ccatatctaa 540

```



```

atgtagaaaa ggtacagtaa aaatatggta taatcttatg ggatcaccat catatatgca 600
atcctttgta gactgaaatg tcattgtgta gtgcatgact gtatacgcac acatacacaa 660
acacacacaa atatactatt ggttcttttt ctctgaagag ccctaataca atagtgtata 720
catttatatt gactctattt caaaatttat ggttttggtg aaacatatgt ggagatgggg 780
catagggtgtg tgaactggga tagtgtctcg ctgatgaatg ggtgggaggc atcatttggg 840
acaagcccag ggcatcagct tatagatatc aagagctcaa caagagcact ttatggcaaa 900
acctcccaca agacctctca gaagttgaga aactgctaaa agtttcttta tgacagatga 960
catttatgga taaaataggg attagcagga ttctttaaat actttcgaac actaaccttc 1020
atttctacca ggcagtgggg cccaagtgc agggccatag gaagtacaag tctgggagat 1080
actaggctgc actgtctgta gagaatctga aaaaataata gagtcaactga aatgcagttt 1140
gggtataatta ttgccatgca tcataattct aatcatact agtgggtcaaa tactcttccc 1200
tgaaaaaaca ttttcttggt ttgaaattcta aataattggt gtgggtacca ctgagctttt 1260
aaatatataa atactttcaa gtttgcata ttttattacc tgttctctaa caaacattga 1320
attcaacatg aaaaatgatta tgggaaacat tcgggtatac agtccctgac tcttaaggac 1380
tcaggtaaat acttagggta tttcatggcc ctagtctttg gggtaggaca tgttcttct 1440
tcaaatcaca gattcaaaat caagaatgat aaacagtgta ttgtgtagac aaaataagt 1500
aaccaaaatt gcttgcctct gtcattctat ggaaccactg agagttttta cttgtgctta 1560
aaattttgaa tagttaaaca gaagtgtcaac ttcattgtgg aatatttttg gctttttaga 1620
cacaatttta agtacctgaa gtatttttac aagactaagt aacatcactg aaattacagc 1680
tttcttcttt ttaaaactgg tatttggtat aaaactaaag agcgaatcaa gaaaagcata 1740
attattactg attattacag gattattact gaaaaagaaa tgtacggaat agaggaggaa 1800
ggagttaaca aatgatccac tctgggtggt gaaaacacca ataagcctgc ttccagggaag 1860
tgcctaagac agagctggct cagcttgctg ggtcacagca tgtaaggaaa ctgctgggct 1920
acatgccacc atctcagtt gtccagatag ataatcccat agccccatgg ggaaataatc 1980
tttaattatg atatagctga caccattcaa agcactatgc taagtccttt atgtgaatta 2040
acttttgtca aatttatttt tcataaataa cccaatatg tataccacta ttatcctacc 2100
ttaagagga gaaactgagc tcctaaagtt taaatatcta acccaagtta agactgctag 2160
tcaccctagg ctattaactc aggcagtcta actcaggtat aataacatta tgctactgtt 2220
tgcagctttg actatgcctg aattataacg tcatgctatc taactaaaaa gctaaggga 2280
ataaaatgag ccatagggtt caatttcata aaaggagaga aaatactggg gaaaagtgat 2340
aatgcagagt ttaaaatatt tttgtaaaag tgccagagat tgagtataac aagtgtgacc 2400
aaaaaataaa aaaaaaagaa aaaagggaag aggtaaaaaa aagagggagg tctgagaaat 2460
agaaatatca gaggaaggaa ataaaggagg gtgagagtaa attctctttt agcattcaga 2520
ttccacagat tccacaatc acatttcttt ttttaccac taaggaaaaa taacacttga 2580
cctaacattt cattgcagtt agctaaagga tgctagaaaa actatgttgc agtggtttgc 2640
tctaatttct tcaggaatag agaaaagtga caaaaagatc agagaagaga agaaaggaaa 2700
ctatcagaaa aatacagaat tggagtagga tataacatat ttgggttgaa ggtaaaattt 2760
tatattgtaa tcttaagtat ctgtctactt cagtttggtc cctggaacag cagcatcaga 2820
atctgccgag ggcttggtta aaaggcagaa tctcaggtcc catcccagac tcaactgaatc 2880
agaatataaa tactgacaag atgccccggg attcatatgc acagtagagc tggcgaagtt 2940
ccattgtagc ctgtgattgt tttctgcaac ttagtatttc tgagttttcc caaggaagaa 3000
aaccagggcc ttagcttctg gcagacttgt gtttctcctt tacttactag ctgcatgact 3060
catgagcaag gaaatcaaac tttatgtgcc tgagtttctt catctataaa atggagacta 3120
taataatcat ctctaggct tgttttgagg atgttcaaca aatgctcctt tcattcctct 3180
atctcagac ctgccgcaga caattctgct agcagccttt gtgctattat ctgttttcta 3240
aacttagtaa ttgagtgtga tctggagact aactctgaaa taaataagct gattatttat 3300
ttattttctc aaaaacacag aatacgattt agcaaattac ttcttaagat attattttac 3360
atctctatat tctcctaccc tgagttgatg tgtgagcaat atgtcacttt cataaagcca 3420
ggtatacatt atggacaggt aagtaaaaaa catattattt attctacgtt tttgtccaaa 3480
aattttaaat tcaactgtt gcgcgtgtgt tggtaatgta aaacaaactc agtacagtag 3540
tattcagtag agtattttaag cccctgtact taaacatatt cctcgtagca atgaagttac 3600
atgaaaagca aatttggtg agatatcgta gatggaagta aattagtctt tatgttcccc 3660
acaaattgaa atgcatttca aaaactctgt gtgtgtatgt gtgtgtgtga cagagtgtgt 3720
gtgagagaga gacagagaga tacgcttttg ttgcctccat aagctggctg ctatgattaa 3780
taagaccaag ttttctaaag aaaatgagat catacaaaaa gccctcttta tgactatctt 3840
ttatcagggg caaaaaggaa agagacaaaa cagcatgaaa tgatgagacc aagtgatgaa 3900
aatctattca caatgattgc tttcctctt aaattctctt gggtaattca gcagcctgtt 3960
actatggctc tctggagtga tagctaagt aatgaagcc tctaaaagtg gattatcctg 4020
acaagaatat actcagccaa taatgcaaca gaaatccatt caaagcattc gggaaaaatt 4080
caaaagaata aatattcttt tttttttttt aaagttaatg acctacgatc catttcttcc 4140
ctgactaaca agcagcaagc acttaaaaaa atccagccag gatgaaatag aaaccacact 4200

```

gacttggttaa	tatttttgg	tggtcccagg	gactcagatt	ctaagccaaa	ttctttgaat	4260
gatcttggca	aatgtctcga	attattttttg	ccaacttttc	tttatcttgg	aaaaaaagtt	4320
tcatgaatgg	gtgtcaaaaat	tgattagttt	taaaaacctt	tcttgcagat	acgtatggca	4380
ccctaaaact	gtattagaaa	aaaagtaagt	actctgtagt	gtgaaaaatt	cttaaaggac	4440
acctctttt	acaaactcac	aaaaacagcc	tttggaaatc	ccacatgaag	tagctgttgt	4500
tattgctttc	tatataccta	catcttgtct	attataaaaa	gactggtttt	tggcaggtgt	4560
ggtggctcac	acctgtaatt	ccagcacttt	gggaggccaa	ggcgggcgga	tcacctgaga	4620
tcaggagttc	aggaccagcc	tgatcaatat	ggtgaaaccc	agtctttact	gaaaatacaa	4680
aaatcacccg	ggtgtggtga	cgggcgcctg	tagtcccagc	tactcgggta	gctgaggcag	4740
gagaatcact	tgaactcagg	agtcagaggt	tgcagtgagc	tgagatcatg	ccactgcact	4800
ccagcctggg	tgacagagca	agactccatc	tcaaaaaaaa	aaaaaaaaaa	gactggtttt	4860
tcaacagctt	tccccacccc	tctgcattga	aatattcacc	cagtcaattg	ttttcctagt	4920
ttgggtaatg	gccctctggg	caggactgga	gtggggcaca	caggagaagc	tgcaaaactat	4980
gtttagaagc	atgtctggga	aatgtcatgc	aagaaaagac	atattttaaag	gtaggctttg	5040
catgaatgga	aaaggagagt	aattctatgt	agagcagagc	ctcttacttg	cagtgaagaga	5100
agcaaaagtg	gggaagcaag	aggaattatg	cttttcatca	gccaaatttg	caggtaggag	5160
gattggctca	gtcatcttgg	ctgaggctca	tgaaccagg	tgtaaagaaa	gtggactaga	5220
ttaatttcat	ccattacagg	aagaggagcc	gtgaaagata	atccagaaat	cattgggatt	5280
tgatggtaga	aggtattttg	ggactatttc	atttgaatg	agaaggatcc	tgacattctt	5340
tgaattcctt	tcaagcaaaag	gattaaattt	acccatgagt	tgactcagaa	aaaacataaa	5400
aagtattgtt	gctctgctca	gagttttatc	taactcattc	tcacttctta	ttccatgatg	5460
aaatgacata	aatgagggtt	tttattgttg	ttgttgttgt	tttctggaca	caaggcaagg	5520
tagctacctg	ggcagagctg	ttttatttct	ctatgccgtg	gagagaaatt	ggttaattgg	5580
ccatggaagg	cagtcattaa	gatgttccca	tgcgagtga	ctttccaggg	ttcccagctt	5640
ctgcatcctt	cctgtccct	caattccatt	gttggtagat	acaatgtctc	tcccatcage	5700
ctcatggaagt	tctctctcat	ttattaaaat	ttgctttcag	gaaaaatttt	gaaaatgtgt	5760
ccagtaatgc	ctgattggcc	ccttatccta	aaggcttaaa	ctggagggaag	gaagctaaac	5820
tgagaaatct	tgcaaatcat	tgagccaaaa	acgtattaat	agcaagatct	atcattttatt	5880
gactagtatg	tggcaggcag	tgccctttta	tttaggcagg	gagagttgat	ggggggggcg	5940
gggttcacac	atcttaaaga	ggtgctatct	cctcctatat	aatcatgta	agtcaagaga	6000
gtaaggaatt	gtctttgttt	ggttatattc	aggggattag	agtatacagt	agaagatccc	6060
aagaaacctt	gggatcattt	tagactaaga	aatgccaaata	ccgccggcg	cggtaggctca	6120
cgctgtaat	ccagcactt	tgagaggccg	aggtgggcgg	atcacaaggt	caggagattg	6180
agaccgtcct	ggctaacgtg	gtgaaaccct	gtctctacta	aaaatacaaa	aaattagccg	6240
ggcgtgggtg	cgggcgcctg	tagtcccagc	tactcgggag	gcggaggcag	gagaatgggtg	6300
tgaactcagg	aggcggagct	tgagtcagc	cgagattgcc	ccaatgcact	ccagcctggg	6360
cgacagaacg	agactccgtc	tcagaacaaa	acaaaaggaa	atgccaatac	cagcagaaat	6420
agagccaaat	actgaacata	agctaaacaa	atgttggcag	tgtagcctag	tggttaagag	6480
agcagactct	taactagaac	actgcactcc	atgtcctcac	tgtagaccct	cactgtgggg	6540
ttctaattaa	cccctgttac	ttaccagtgg	cagtcttaag	gcattcctta	agttcgttgt	6600
gccccaaatt	gttcatctgt	agaaggggta	ggatgacagt	agtgtttact	ttataggctt	6660
actgtgagca	ttaaatgagt	tactactgta	tttgtaaagt	gcttaaaatg	ctgctccaaa	6720
agagtttgtt	aaacacttaa	gaactgattt	acttgcactc	aaactgacag	ctctcaataa	6780
ctggaatatg	tcaagcatag	gccctggaat	ataagcagg	ctacatgaag	gcaaaaatgt	6840
tcgtttcttt	tgttcagccc	tgtgcctaga	tcaatatcta	tgatcatgc	tcaagaaata	6900
ttgttgaatg	aatcaatgaa	cctaccgagg	tagttacata	aaagagttct	gcatgagtag	6960
aaatctgggc	aaagtgcct	ccaaggaaat	ttccactttt	agattctgtg	atttccttaa	7020
ggaactgata	aattggtgtg	atacaatgta	aaaaaatgtg	cctatatgat	ttgagaaaaa	7080
cttattttct	ctccctcttt	tttcttctct	tccttctctc	ctcccttctc	tccttctctc	7140
ctcccttctc	tcctccctcc	ctcccttctc	tccttctctt	tcttcttttc	tttctttctt	7200
tctttctttc	tttctttctt	tctttctttc	tttctttctt	ctcttctctt	tcctttcttt	7260
cttctttctt	ttgtgccttt	cttctttctt	tttctttctt	tctgtccttt	cttcttctct	7320
ttctttcttt	ctgcctttct	ttctttttgt	tttctttctt	ttccttcttt	tttcttttaa	7380
gcagaccatg	tctgttagat	gaatgccttt	ttctagttaa	aagggttaa	aggaaagtga	7440
agcacaatta	tcaagggctc	ccagtcactc	ccacatgttc	ttaatcatta	tcttctttta	7500
cagtttcata	tctccaggcc	tttcatggg	tcaggttggc	atttcgctgc	cctttatgtg	7560
tgtgacaagt	gaaaataagg	aaagaaaaaa	actcaagtga	agaaaatcag	aatctgcgca	7620
gcagttcctg	ggcgtttcag	ctgcttccca	catcacctgc	ctcatcaagc	cccagcatcc	7680
atctccttgc	tcatcttaca	cctgtgtgtc	atgacaggcc	caccattcat	ttatcagagc	7740
aaaggctctc	ccactattct	ggttcacccc	cctacttagc	cagatatata	agaatatctg	7800
cacggatgac	ctgcctcacc	tgggagctca	gaggagctca	gattccatta	ctatcgacc	7860

```

aaggacagat ctcccagcaa gaatgacaga aaagactaac tgcccccaa atctcccttc 7920
caaaacacag ttctcttaat tctcccaaga aaccagaatg tgactgctca cctctctaag 7980
gacctgaaaa caactggcca tttagctat ttaaataaac tttaaaaaat ccaaccgcca 8040
aaatattaaa ccattttggt tggaatgata acataactaa cctgctgaca gctgcttctg 8100
ctaggtgcaa aaatggaaaa aaaaatactt ctaatcaggt caaatcactc tacctttggg 8160
attctaaatt tactcatatt ctcaaagaaa tatattcagt catagtggg aaaataggat 8220
tattccttta gctcgataag caaccagaag ttcttccttc aaatcttgac atttaatcaa 8280
tcagaaattg atttttggaa aactgtttcc tatgaagcta tctctgcctg aaggattttt 8340
cttttacaat ccagactata gaaggaaatt cacaacctgg actttcacct ccattggtca 8400
gagttttact gaccaattcc cacctctgcc ttacacctaa cggaagtta tgctgtttt 8460
ctcttcacat accccaacag ttacaaatgg ttgttattat taagcatctt ttattttgtg 8520
gcctctgatt acatgggtccc cttaaattttg acctaatacac aaaagattgg taaaatttct 8580
taacatatta ataataattt gtttatgtgt caatatctta gcatgtatca attaagacag 8640
aggtcttaac gttctctttt tgaaagagaa tattaggatt cagagatatt aagagattct 8700
cccaggatca cagttaggta acagagctgg attttagtcc aggtctgtct acagctctaa 8760
cgtatataca ccctttgtat aacatgtcac gaattcagca taaagggatc ttcagtgatc 8820
taagtcaggg gtcagcaacc ttttctaaaa aggaccaaat agtaatatct caggctttgt 8880
ggaccctatg gtctctatca taactgttca aatcaccatg tagtgtaaaa ggagccataa 8940
gcaaaatata aactaacgaa tgtggctgtt ttatgggatt tttttttaac tctttattta 9000
caaaaagcagg tggcagatca gaactcactt atgggccata gttctctgac ccctgacctg 9060
agaaaatctt atatttatgg acaacattta gactgtgact tgccaagtaa gaacaagaag 9120
ctctgtcaac tgaagggtcaa ggctggagtt ctgaaagcaa agagctgtct ggtgttaatg 9180
ataagtgaat tagttaaagt tagaagatcc cagttataag aagcacaaaag aataatgacc 9240
atagactcct gaacaagaat gtctggactt ctggcttagg cactcttgtt gtatggtcca 9300
ggccaagtta cctaactctt ccaggcctcc attttcttat cattaaatga agataataaa 9360
agtattttcc tcagagagct gtaagaataa actgagctaa cccatgtcaa gcacatagaa 9420
tagggcccag cctatattaa tttatcaata aatgccag 9458

```

<210> 12
 <211> 552
 <212> DNA
 <213> Homo. sapiens

<220>
 <221> exon
 <222> (0)...(0)
 <223> Exon 1

```

<400> 12
attatgtaga accattccaa tcctgagtga caggatttgc aaagacagct gttgggctcc 60
acaaacattt ggaggagggg aaaatgtgag gttggttgct acattctcct acttcttttt 120
actgaaagga acagctgcga tcttcacatg taagatgaag taaacaaaag ctgacaatgc 180
ccagcaccat tcagcagtaa gcattcaaag ggatttttcc tcaccgcgtt tgcaggcccg 240
gcacagtctc agttccccac agcacagaga ggaagcaagg tctttgcatg ccctgttcac 300
ttggccaaac attgccttag atgtgcaatt caacccatat aaagcaagcg taattcaaga 360
gccctgactc tttcaccatt tcatgatggc ttattactcg cacacggcat tgtcagggtc 420
cagcaactct gtgtgcagca ggcagaaatg cagtgtttat ctctgacagg caaacattct 480
taggctcatc ggcatgagta tccacgctga actggctatc ggggaagtag atccctgcat 540
gggaaagtca ag 552

```

<210> 13
 <211> 305
 <212> DNA
 <213> Homo sapiens

<220>
 <221> exon
 <222> (0)...(0)
 <223> Exon 1a

<400> 13
atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcacatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaagga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttacac ctttggtctc cctgggtctc agttccttgg acttggaactg aattacactg 240
ccagggtttcc tgggttctcca gcttgcagat ggcagatcat gggacttctt ggcctccata 300
attgt 305

<210> 14
<211> 307
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (0)...(0)
<223> Exon 1b

<400> 14
atttgaattg gtgaacttag taaagcagac ggctctcacc aataagggca ggcacatcc 60
aatctgtcga aagcttgaat aaaacaaaaa gaggaagga aaatttgctt cttttcttct 120
tgatctagta tatcatcttc tcctgccctt ggatgtgagt gggccttcag acttaaacca 180
ggagttacac ctttggtctc cctgggtctc agttccttgg acttggaactg aattacactg 240
ccagggtttcc tgggttctcca gcttgcagat ggcagatcat gggacttctt ggcctccata 300
attgtgt 307

<210> 15
<211> 557
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (0)...(0)
<223> Exon 2

<400> 15
gggcaaaaag gaaagagaca aaacagcatg aaatgatgag accaagtgat gaaaattcat 60
tcacaatgat tgctttcaag agtaatttct cttgggtaat tcagcagcct gttactatgg 120
ctctctggag tgatagctaa tgtaaatgaa gcctctaaaa gtggattatc ctgacaagaa 180
tatactcagc caataatgca acagaaatcc attcaaagca ttcgggaaaa attcaaaaaga 240
ataaatattc tttttttttt tttaaagtta atgacctacg atccatttct tccctgacta 300
acaagcagca agcacttaaa aatatccagc caggatgaaa tagaaaccca cctgacttgt 360
taatattttt gtttggtccc agggactcag attctaagcc aaattccttg aatgatcttg 420
gcaaatgtct cgaattattt ttgccaaactt ttctttatct tggaaaaaaa gtttcatgaa 480
tggtgtgtaa aattgattag ttttaaaaac ctttcttgca gatacgatg gcaccctaaa 540
actgtattag aaaaaaa 557

<210> 16
<211> 1955
<212> DNA
<213> Homo Sapiens

<220>
<221> exon
<222> (0)...(0)
<223> Exon 3

<400> 16
tttcatatct ccaggccttt cattgggtca ggttggcatt tcgctgccct ttatgtgtgt 60

```

gacaagtgaa aataaggaaa gaaaaaaact caagtgaaga aaatcagaat ctgcgagca 120
gttcctgggc gtttcagctg cttcccacat cacctgcctc atcaagcccc agcatccatc 180
tccttgctca tcttacaccc tgtgtgcatg acaggcccac cattcattta tcagagcaaa 240
ggctctccca ctattctggg tcacccccct acttagccag atatacaaga atatctgcac 300
ggatgacctg cctcacctgg gagctcagag gagctcagat tccattacta tcgcaccaag 360
gacagatctc ccagcaagaa tgacagaaaa gactaactgc ccccaaaatc tcccttccaa 420
aacacagttc tcttaattct cccaagaaac cagaatgtga ctgctcacct ctctaaggac 480
ctgaaaacaa ctggccattt cagctattta aatcaacttt aaaaaatcca accgccaaaa 540
tattaaacca ttttggttgg aatgataaca taactaacct gctgacagct gcttctgcta 600
ggtgcaaaaa tggaaaaaaa aatacttcta atcagggtcaa atcactctac ctttgggatt 660
ctaaatttac tcatattctc aaagaaatat attcagtcac agtggggaaa ataggattat 720
tcctttagct cgataagaa ccagaagttc tctcttcaaa tcttgacatt taatcaatca 780
gaaattgatt tttggaaaac tgtttcctat gaagctatct ctgcctgaag gatttttctt 840
ttacaatcca gactatagaa ggaaattcac aacctggact ttcacctcca ttggtcagag 900
ttttactgac caattcccac ctctgcctta cacctaacgg aagtttatgc ctgttttctc 960
ttcacatacc ccaacagtta caaatgggtg ttattattaa gcatctttta ttttgtggcc 1020
tctgattaca tgggtccccta aattttgacc taatcacaaa agatttgtaa aatttcttaa 1080
catattaata atattttgtt tatgtgtcaa tatcttagca tgtatcaatt aagacagagg 1140
tcttaacggt ctctttttga aagagaatat taggattcag agatattaag agattctccc 1200
aggatcacag ttaggtaaca gagctggatt ttagtcagg tctgtctaca gctctaactg 1260
atatacacc tttgtataac atgtcacgaa ttcagcataa agggatcttc agtgatctaa 1320
gtcaggggtc agcaaccttt tctaaaaagg accaaatagt aatatttcag gctttgtgga 1380
ccctatggtc tctatcataa ctgttcaaat caccatgtag tgtaaaagga gccataagca 1440
aaatataaac taacgaatgt ggctgtttta tgggattttt ttttaactct ttatttacia 1500
aagcaggtgg cagatcagaa ctcaattatg ggccatagtt ctctgacccc tgacctgaga 1560
aaatcttata tttatggaca acatttagac tgtgacttgc caagtaagaa caagaagctc 1620
tgtcaactga aggtcaaggc tggagttctg aaagcaaaga gctgtctggt gttaatgata 1680
agtgaatatag ttaaagttag aagatcccag ttataagaag cacaagaat aatgaccata 1740
gactcctgaa caagaatgtc tggacttctg gcttaggcac tcttgttgta tgggtccaggc 1800
caagttacct aatctctcca ggctccatt ttcttatcat taaatgaaga taataaaagt 1860
attttctctc gagagctgta agaataaact gagctaacc atgtcaagca catagaatag 1920
ggcccagcct atattaattt atcaataaat gccag 1955

```

<210> 17

<211> 366

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (0)...(0)

<400> 17

```

atgttcaaca aatgctcctt tcattcctct atttacagac ctgccgcaga caattctgct 60
agcagccttt gtgctattat ctgttttcta aacttagtaa ttgagtgtga tctggagact 120
aactctgaaa taaataagct gattatttat ttattttctc aaaacaacag aatacgattt 180
agcaaattac ttcttaagat attattttac atttctatat tctcctaccc tgagttgatg 240
tgtgagcaat atgtcacttt cataaagcca ggtatacatt atggacaggt aagtaaaaaa 300
catattattt attctacggt tttgtccaaa aatttttaaa ttcaactgtt gcgcgtgtgt 360
tggttaa 366

```

<210> 18

<211> 1253

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> (0)...(0)

<223> ESTbc03936 Exon 2

```

<400> 18
tatccccact gtgtagagct ggtccacaat ggagaggagg aaaagttttg cattggacca 60
aatcactgcc ttagaaggaa agagattgca atcccattca tttttaagat ctgcaatggg 120
ttgggatgat acccatttaa aaaagaaaaa aaaagagggtt acctctttgg cctctttggc 180
aaaggatatg aactagtaat aaagcctcag catcaaggct tgggagtctg gctcaaagtg 240
gctaagagcta gaagcatctt ctaacataaa tatggcagca ttaggaaaat agcttggcct 300
ttagagagac ccagattatc ttcaccttaa aaactaccaa gaccctggtc cagcaagacc 360
accatcacag cagtaaatac atagccacca acttcacaga aggttttcca ctgagacatt 420
gtgaaaacca caatocatta ccggtggatg agtcccagct ctcttactac ctggagcttc 480
caagcaagac ttacctgtcc tttgttccag attggttcct gggatcctgg accaacacaa 540
agagaagagg aaagcctcaa gggaaataaa agtgaagggt tcagtgatag tttaaatata 600
tattttttaa tgtagctttt tttttaataa cttcacgggt atggatggag ttttcaacc 660
acaatggcaa agaatgaatt accggcagtg tcagagtttc ttgttgctgt gggaagacag 720
ctgccatgtt gagagggtgcc ttgtggagag gctcatgttg cgaggaactg agactggctt 780
ctatccaaca accagtggag aactaaggcc ctaaagtcct atagtccaca aggaactaaa 840
tcctaccaac aactacatga gtgagcctgg aatctgattt tcccagtggt ggacttaagg 900
tacctacaac ctggctgtta tgggctaaac tgcctcctgc ccaaatttat atgtttgagt 960
cttaactccc agcacttcag aacatgggtt tgtttggaga catagtcatt aaagaggtaa 1020
tgaagttaaa atgagggtcat taggatgagt cctaattcaa taggactggt gttcttatat 1080
gaagaggaaa tttgaacaca gacacatata caggaaagac catgcgaaga agacacagag 1140
aggaaacggc catctacaat ccaaggagag aggcctcaga acaagccaac cctgcagata 1200
ccttgatcta ggcattccaga attgtatgaa aatacathtt tgttatttaa gct 1253

```

```

<210> 19
<211> 21
<212> RNA
<213> Homo sapiens

```

```

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 1

```

```

<400> 19
aauaagggca ggcaucaucc a 21

```

```

<210> 20
<211> 21
<212> RNA
<213> Homo sapiens

```

```

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 1 B

```

```

<400> 20
aauuacacug ccagguuucc u 21

```

```

<210> 21
<211> 21
<212> RNA
<213> Homo sapiens

```

```

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 2 A

```

```

<400> 21
aauucauuc caaugauugc u 21

```

<210> 22
<211> 21
<212> RNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 2 B

<400> 22
aauuucucuu ggguuuu g 21

<210> 23
<211> 21
<212> RNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 3 (DDend) A

<400> 23
aaaucagaa ucugcgcagc a 21

<210> 24
<211> 21
<212> RNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting Exon 3 (DDend) B

<400> 24
aaugaugaug ggaagaagga a 21

<210> 25
<211> 21
<212> RNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (0)...(0)
<223> siRNA targeting CDS A

<400> 25
aaacuagua auugagugug a 21

<210> 26
<211> 21
<212> RNA
<213> Homo sapiens

<220>
<221> mRNA

<222> (0)...(0)
<223> siRNA targeting CDS B

<400> 26
aauaugucac uuucauaaag c 21

<210> 27
<211> 14
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> Exon overlapping oligonucleotide primer F26

<400> 27
agacggctct cacc 14

<210> 28
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> Exon overlapping oligonucleotide primer FDP5

<400> 28
cacttgtcac acacataaag g 21

<210> 29
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> cDNA oligonucleotide primer FDP5

<400> 29
cctttatgtg tgtgacaagt g 21

<210> 30
<211> 22
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> cDNA oligonucleotide primer F10

<400> 30
atccagccag gatgaaatag aa 22

<210> 31
<211> 20
<212> DNA

<213> Homo sapiens

<220>

<221> primer_bind

<222> (0)...(0)

<223> upstream primer human VH family leader sequence
VH1

<400> 31

ccatggactg gacctggagg

20

<210> 32

<211> 21

<212> DNA

<213> Homo sapiens

<220>

<221> primer_bind

<222> (0)...(0)

<223> upstream primer human VH family leader sequence
VH2

<400> 32

atggacatac tttgttccag c

21

<210> 33

<211> 20

<212> DNA

<213> Homo sapiens

<220>

<221> primer_bind

<222> (0)...(0)

<223> upstream primer human VH family leader sequence
VH3

<400> 33

ccatggagtt tgggctgagc

20

<210> 34

<211> 20

<212> DNA

<213> Homo sapiens

<220>

<221> primer_bind

<222> (0)...(0)

<223> upstream primer human VH family leader sequence
VH4

<400> 34

atgaaacacc tgtggttctt

20

<210> 35

<211> 20

<212> DNA

<213> Homo sapiens

<220>

<221> primer_bind

<222> (0)...(0)

<223> upstream primer human VH family leader sequence
VH5

<400> 35
atgggggtcaa ccgcgatacct 20

<210> 36
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> upstream primer human VH family leader sequence
VH6

<400> 36
atgtctgtct ccttcctcat 20

<210> 37
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> down stream primer C1

<400> 37
gaggctcagc gggaagacct t 21

<210> 38
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> down stream primer C 2

<400> 38
ggggaagacc gatgggcccc t 21

<210> 39
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> northern blot primer

<400> 39
tcacctggga gctcagagga 20

<210> 40
<211> 20

<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> northern blot primer

<400> 40
gtgatcctgg gagaatctct

20

<210> 41
<211> 25
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (0)...(0)
<223> RACE antisense primer

<400> 41
tacattacca acacacgcgc aacag

25